

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:51:09 ; Search time 50.1316 Seconds
(without alignments)
2147.361 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALAVVLLHLTR.....ANEAAPFRLFNHDKFRD 381

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	381	2	Aaw35730 Human Cys
2	2116	100.0	381	4	Aae05921 Human Cys
3	2106	99.5	381	2	Aaw35957 Human mon
4	2106	99.5	381	4	Aab30773 Human she
5	2106	99.5	381	5	Aau79761 Human Cyr
6	2106	99.5	381	5	Abb05438 Human Cyr
7	2106	99.5	381	5	Aae18107 Human con
8	2106	99.5	381	6	Abu63222 Human can
9	2106	99.5	455	3	Aab43987 Human can
10	2098	99.1	381	5	Abg76937 Human pro
11	2098	99.1	381	7	Abb75260 Prostate
12	1971.5	93.2	374	5	Abb09202 HCGF CNN
13	1971.5	93.2	375	2	Aar90919 Connectiv
14	1971.5	93.2	375	2	Aay31620 Human CTG
15	1971.5	93.2	375	5	Aae18108 Human alt
16	1971.5	93.2	375	6	Abu63223 Human con
17	1929	91.2	379	2	Aar25565 Beta-IG-M
18	1929	91.2	379	4	Aae05920 Mouse cys
19	1929	91.2	375	5	Abb09201 Mouse Cyr
20	1695	80.1	375	4	Aae05939 Chicken C
21	1695	80.1	375	5	Abb09203 Chicken c
22	955	45.1	347	6	Abra43138 Rat conne
23	953	45.0	347	2	Aay24379 Rat conne
24	952.5	45.0	348	2	Aar25566 Beta-IG-M
25	952.5	45.0	348	6	Adb25766 Mouse con

ALIGNMENTS

RESULT 1
Aaw35730
ID Aaw35730 standard; protein; 381 AA.

XX Aaw35730;
XX AC Aaw35730;
XX DT 27-MAR-1998 (first entry)
XX DE Human cysteine rich protein 61 (Cyr61).
XX KW Cysteine rich protein 61; Cyr61; human;
XX KW extracellular matrix signalling molecule; cell adhesion; cell migration;
XX KW cell proliferation; angiogenesis; chondrogenesis; oncogenesis;
XX KW haematocytosis; wound healing; organ regeneration.

OS Homo sapiens.

XX WO9733995-A2.

PD 18-SEP-1997.

XX PF 14-MAR-1997; 97WO-US004193.

XX PR 15-MAR-1996; 96US-0013958P.

XX PA (MUNI-) MUNIN CORP.

XX PI Lau LF;

XX DR WPI; 1997-470875/43.

XX DR N-PEDB; AAT94699.

XX PT Isolated and purified cysteine rich protein 61, Cyr61 - useful to
modulate e.g. haematocytosis, induce wound healing, promote organ
regeneration etc.

XX PS Claim 2; Page 112-113; 133pp; English.

XX CC This protein sequence comprises human cysteine rich protein 61 (Cyr61),
an extracellular matrix signalling molecule. Its amino acid sequence was
deduced from a human placental cDNA clone (see AAT94699). Cyr61
polypeptides can be expressed in transformed or transfected host cells.
Cyr61 can be used to modulate haematocytosis, induce wound healing in a
CC tissue, promote organ regeneration, improve tissue grafting or promote
bone or prothesis implantation (claimed). It can also be used to screen
CC for a modulator of angiogenesis, chondrogenesis, oncogenesis, cell
adhesion, cell migration, cell proliferation, expand a population of
CC undifferentiated haematopoietic stem cells in culture and to screen for a

Adb25762 Mouse con
Aaw35731 Murine Fi
Aay44756 Mouse con
Aae05922 Mouse fib
Abb09205 Fisp-12 C
Abr43139 Mouse con
Aaw12694 Connectiv
Aay93340 Amino aci
Aar79964 Connectiv
Aaw11302 Connectiv
Aaw09089 Human con
Aaw2084 Human con
Aaw81425 Connectiv
Aay18361 Human con
Aay92939 Human con
Aay92940 Human con
Aay44755 Human con
Aab84598 Amino aci
Aab90791 Human she
Aab60664 Human con

26 952.5 45.0 348 6 ADB25762
27 948.5 44.8 348 2 AAW35731
28 948.5 44.8 348 3 AAY44756
29 948.5 44.8 348 5 AAE05922
30 948.5 44.8 348 5 ABB09205
31 948.5 44.8 348 6 ABR43139
32 948 44.8 347 2 AAW12694
33 948 44.8 347 3 AAY93340
34 947 44.8 349 2 AAR79964
35 947 44.8 349 2 AAW11302
36 947 44.8 349 2 AAW09089
37 947 44.8 349 2 AAW2084
38 947 44.8 349 2 AAW81425
39 947 44.8 349 2 AAY18361
40 947 44.8 349 3 AAY92939
41 947 44.8 349 3 AAY92940
42 947 44.8 349 3 AAY44755
43 947 44.8 349 4 AAB84598
44 947 44.8 349 4 AAB90791
45 947 44.8 349 4 AAB60664

CC mitogen (claimed). Ex vivo methods for using mammalian extracellular
CC matrix signalling molecules to prepare blood products are also provided
XX
SQ Sequence 381 AA;

Query Match 100.0%; Score 2116; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 8.9e-162;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
Db 1 MSSRIARALAVVTLHLTLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVTGQCCBEWVDEDSIKDPMEDQDGLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVTGQCCBEWVDEDSIKDPMEDQDGLG 180

QY 181 KELGFDASEVELTRNNELIANGKRSIKRLPVFGMEPRILYNPLQSKCIVQTTWSQCS 240
Db 181 KELGFDASEVELTRNNELIANGKRSIKRLPVFGMEPRILYNPLQSKCIVQTTWSQCS 240

QY 241 KTCGTGISTRVTNDNPECLVKETRICVPRPCGQPVYSSIKGKCKSKTKKSPPEPVFTY 300
Db 241 KTCGTGISTRVTNDNPECLVKETRICVPRPCGQPVYSSIKGKCKSKTKKSPPEPVFTY 300

QY 301 AGCLSVKXKYPKYGCGVDCRCCTPQLTRTVKMRFRCEDETFSKNWMIIQSKCNYNCP 360
Db 301 AGCLSVKXKYPKYGCGVDCRCCTPQLTRTVKMRFRCEDETFSKNWMIIQSKCNYNCP 360

QY 361 HANEAAFPFVRLFNDIHKFRD 381
Db 361 HANEAAFPFVRLFNDIHKFRD 381

RESULT 2
AAE05921
ID AAE05921 standard; protein; 381 AA.
AC AAE05921;
XX
XX
XX 24-SEP-2001 (first entry)
XX Human cysteine-rich protein (Cyr61).
XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
KW fibroblast secreted protein; Fibp12; connective tissue growth factor;
KW CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
KW angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
KW heart disease; fibrosis; gene therapy; human.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 163..229
FT /note= "Cysteine free region"
FT Domain 212..281
FT /note= "Domain III"
FT Domain 282..381
FT /note= "Domain IV"
XX
XX WO200155210-A2.
XX
XX 02-AUG-2001.
XX
XX 31-JAN-2001; 2001WO-US003267.
XX
XX 31-JAN-2000; 2000US-00495448.
XX 15-MAY-2000; 2000US-0204364P.
XX 06-OCT-2000; 2000US-0238705P.

XX (MUNI-) MUNIN CORP.
XX Lau LF, Yeung C, Greenspan JA;
XX
XX WPI; 2001-465561/50.
DR N-PSDB; AAD11221.
XX
PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
PT for screening for modulators of cell adhesion, fibroblast cell
XX proliferation, angiogenesis and cell migration.
PS Claim 30; Page 171-172; 186pp; English.
XX
CC The invention relates to extracellular matrix (ECM) signalling molecules
CC involved in cellular response to growth factors. More particularly the
CC invention is directed to cysteine-rich protein (Cyr61), and Cyr61-related
CC proteins such as fibroblast secreted protein (Fibp12) and connective
CC tissue growth factor (CTGF) and nucleic acid molecules encoding such
CC proteins. The polypeptides of the invention are members of cysteine-rich
CC secreted protein family. Human Cyr61 fragment is useful in methods for
CC screening modulators of cell adhesion, cell migration, fibroblast cell
CC proliferation, angiogenesis, wound healing and Cyr61-integrin receptor
CC interaction. Modulator of Cyr61-integrin alphaVbeta3 interaction is used
CC for the preparation of a medicament for the treatment of atherosclerosis,
CC heart disease, tumour metastasis, fibrosis, tumour growth, disorders
CC associated with inadequate angiogenesis, aberrant granulation tissue
CC development; aberrant fibroblast growth and wounds. Polynucleotides of
CC the invention are useful in gene therapy. The present sequence is human
CC Cyr61 protein
XX
SQ Sequence 381 AA;

Query Match 100.0%; Score 2116; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 8.9e-162;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
Db 1 MSSRIARALAVVTLHLTLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVTGQCCBEWVDEDSIKDPMEDQDGLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVTGQCCBEWVDEDSIKDPMEDQDGLG 180

QY 181 KELGFDASEVELTRNNELIANGKRSIKRLPVFGMEPRILYNPLQSKCIVQTTWSQCS 240
Db 181 KELGFDASEVELTRNNELIANGKRSIKRLPVFGMEPRILYNPLQSKCIVQTTWSQCS 240

QY 241 KTCGTGISTRVTNDNPECLVKETRICVPRPCGQPVYSSIKGKCKSKTKKSPPEPVFTY 300
Db 241 KTCGTGISTRVTNDNPECLVKETRICVPRPCGQPVYSSIKGKCKSKTKKSPPEPVFTY 300

QY 301 AGCLSVKXKYPKYGCGVDCRCCTPQLTRTVKMRFRCEDETFSKNWMIIQSKCNYNCP 360
Db 301 AGCLSVKXKYPKYGCGVDCRCCTPQLTRTVKMRFRCEDETFSKNWMIIQSKCNYNCP 360

QY 361 HANEAAFPFVRLFNDIHKFRD 381
Db 361 HANEAAFPFVRLFNDIHKFRD 381

RESULT 3
AAW35957
ID AAW35957 standard; protein; 381 AA.
XX
XX AAW35957;
XX
XX 05-MAR-1998 (first entry)
DT

XX DE Human monocyte mature differentiation factor.
 XX Human; monocyte; mature; differentiation factor; MMDF; macrophage;
 KW cancer; immune activator; tissue culture; infectious disease.
 XX OS Homo sapiens.
 XX PN JP09234079-A.
 XX PD 09-SEP-1997.
 XX PF 04-MAR-1996; 96JP-00075236.
 XX PR 04-MAR-1996; 96JP-00075236.
 XX PA (TOYM) TOYOBO KK.
 XX PI WPI; 1997-497320/46.
 DR N-PSDB; AAT97142.
 XX A monocyte mature differentiation factor - useful for the long term
 PT tissue culture of macrophage(s).
 XX Claim 9; Page 12-13; 22pp; Japanese.
 XX The present sequence represents a monocyte mature differentiation factor
 CC (MMDF) which maintains the life of macrophages for long periods in liquid
 CC culture. MMDF can be used as an anti-cancer agent, an immune activator
 CC and to treat infectious diseases
 XX Sequence 381 AA;
 SQ
 Query Match 99.5%; Score 2106; DB 2; Length 381;
 Best Local Similarity 99.5%; Pred. No. 5.6e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLCAPKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLCAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDQDGLLG 180
 Db 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTRNNELIYAVGKRSIKELPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
 Db 181 KELGFDASEVELTRNNELIYAVGKRSIKELPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
 QY 241 KTCGTGISTRTVNDNPECLVKEITRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVRFY 300
 Db 241 KTCGTGISTRTVNDNPECLVKEITRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVRFY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPQITRTVQMRFCEDGETFSKNVMMIQSCCKNYNCP 360
 Db 301 AGCLSVKKYRPKYCGSCVDGRCTPQITRTVQMRFCEDGETFSKNVMMIQSCCKNYNCP 360
 QY 361 HANEAAFPYRLFNDIHKFRD 381
 Db 361 HANEAAFPYRLFNDIHKFRD 381
 RESULT 4
 AAB90773
 ID AAB90773 standard; protein; 381 AA.
 XX
 AC AAB90773;
 XX
 DT 15-JUN-2001 (first entry)

XX DE Human shear stress-response protein SEQ ID NO: 46.
 XX Human; shear stress-response protein; vascular disease; arteriosclerosis.
 XX OS Homo sapiens.
 XX PN WO200125427-A1.
 XX PD 12-APR-2001.
 XX PF 02-OCT-2000; 2000WO-JP006840.
 XX PR 01-OCT-1999; 99JP-00280976.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX NOJIMA H. NOJIMA H.
 XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga I, Sekine S, Nakamura Y, Sugano S;
 XX WPI; 2001-266308/27.
 DR N-PSDB; AAH02896.
 XX DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis.
 XX Claim 60; Page 345-346; 678pp; Japanese.
 XX The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension
 XX Sequence 381 AA;
 SQ
 Query Match 99.5%; Score 2106; DB 4; Length 381;
 Best Local Similarity 99.5%; Pred. No. 5.6e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLCAPKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLCAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDQDGLLG 180
 Db 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTRNNELIYAVGKRSIKELPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
 Db 181 KELGFDASEVELTRNNELIYAVGKRSIKELPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
 QY 241 KTCGTGISTRTVNDNPECLVKEITRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVRFY 300
 Db 241 KTCGTGISTRTVNDNPECLVKEITRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVRFY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPQITRTVQMRFCEDGETFSKNVMMIQSCCKNYNCP 360
 Db 301 AGCLSVKKYRPKYCGSCVDGRCTPQITRTVQMRFCEDGETFSKNVMMIQSCCKNYNCP 360
 QY 361 HANEAAFPYRLFNDIHKFRD 381
 Db 361 HANEAAFPYRLFNDIHKFRD 381
 RESULT 5
 AAU79761

ID XX AAU9761 standard; protein; 381 AA.
 AC AAU9761;
 XX 30-JUL-2002 (first entry)
 DT 30-JUL-2002 (first entry)
 XX Human Cyr61 protein.
 DE Human; uterine leiomyoma proliferation; uterine leiomyoma formation;
 XX Cyr61; cytostatic.
 KW Homo sapiens.
 OS WO200226193-A2.
 XX 04-APR-2002.
 PD 28-SEP-2001; 2001WO-US030783.
 PF 29-SEP-2000; 2000US-0236887P.
 XX (AMHP) AMERICAN HOME PROD CORP.
 PA Zhang Z, Sampath D, Zhu Y, Winneker R;
 XX WPI; 2002-383245/41.
 DR N-PSDB; ABK48899.
 XX Preventing uterine leiomyoma formation or inhibiting proliferation of
 PT uterine leiomyoma in subject, comprises modulating or increasing the
 PT level of Cyr61 in leiomyoma tissue.
 XX Disclosure; Fig 6; 92pp; English.
 PS The present invention relates to a method of inhibiting proliferation of
 CC uterine leiomyoma or preventing uterine leiomyoma formation. The method
 CC comprises increasing the level of Cyr61 in leiomyoma tissue. The
 CC invention also describes compounds and compositions that stimulate
 CC induction of the Cyr61 gene and compounds that increase Cyr61 activity.
 CC The compositions and the method of the invention are useful for
 CC preventing uterine leiomyoma formation or inhibiting proliferation of
 CC uterine leiomyoma in a subject. The method is particularly useful for
 CC treating or preventing uterine leiomyoma formation, or inhibiting
 CC proliferation of uterine leiomyoma in a subject. The present sequence
 CC represents human Cyr61. Note: The present sequence shown in Fig 6 is not
 CC shown in the correct sequence order in the figure. The start of the
 CC sequence is shown on page 8/10 of the figures and the rest of the
 CC sequence is shown on page 6/10 of the figures
 XX SQ Sequence 381 AA;
 PS Query Match 99.5%; Score 2106; DB 5; Length 381;
 CC Best Local Similarity 99.5%; Pred. No. 5.6e-161;
 CC Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSRIARALAVTLTLTLALSTCPAACHCPLAPKCAPGVGLVDRDGGCKVCAKQL 60
 DB 1 MSRIARALAVTLTLTLALSTCPAACHCPLAPKCAPGVGLVDRDGGCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCECNFGASTALKGICRAQSGRCPCEYNRIYQNGESFPQNCQHQ 120
 DB 61 NEDCSKTQPCDHTKGLCECNFGASTALKGICRAQSGRCPCEYNRIYQNGESFPQNCQHQ 120
 QY 121 CTCIDGAVGICPLCPQBSLPLNLCNPNRLVKVGTGCCCEWVCDSDIKDPMEDQDGLLG 180
 DB 121 CTCIDGAVGICPLCPQBSLPLNLCNPNRLVKVGTGCCCEWVCDSDIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTNNELIANGKRSILKRLPVFGMEPRILYNPLQCKIVQTSWSQS 240
 DB 181 KELGFDASEVELTNNELIANGKRSILKRLPVFGMEPRILYNPLQCKIVQTSWSQS 240
 QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVRPCQPVYSSILKGGKCKSKTKSPVRFV 300

DB 241 KTCGTGISTRTVNDNPECRIVKTRICEVRPCQPVYSSILKGGKCKSKTKSPVRFV 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPQLTTRTVKRFCEDEGTFPSKVMMIQSKCNYNCP 360
 DB 301 AGCLSVKKYRPKYCGSCVDGRCTPQLTTRTVKRFCEDEGTFPSKVMMIQSKCNYNCP 360
 QY 361 HANEAAFPFVYRLFNDFHKKFED 381
 DB 361 HANEAAFPFVYRLFNDFHKKFED 381
 RESULT 6
 ABB05438
 ID ABB05438 standard; protein; 381 AA.
 XX AC ABB05438;
 XX 15-APR-2002 (first entry)
 DT Human Cyr61 protein SEQ ID NO:2.
 DE Human; Cyr61; breast cancer; sex steroid receptor; cytostatic; promoter;
 XX sex steroid response element; cysteine rich heparin-binding protein;
 KW cell proliferation; heparin binding epidermal growth factor;
 KW epidermal growth factor; basic fibroblastic growth factor.
 XX Homo sapiens.
 OS WO200198359-A2.
 XX 27-DEC-2001.
 PD 21-JUN-2001; 2001WO-US019823.
 PF 21-JUN-2000; 2000US-0213182P.
 PR 16-MAY-2001; 2001US-0291510P.
 XX (AMHP) AMERICAN HOME PROD CORP.
 PA Sampath D, Zhang Z, Winneker R;
 XX WPI; 2002-147796/19.
 DR N-PSDB; ABA93127, ABA93130.
 XX Regulation of Cyr61 expression and activity for preventing and inhibiting
 PT breast cancer comprises use of a Cyr61 neutralizing antibody, an anti-
 PT sense oligonucleotide and an antibody which.
 XX Claim 6; Fig 1; 86pp; English.
 PS The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-
 CC binding protein) promoter. Cyr61 has cytostatic activity. An antibody (I)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal,
 CC heparin binding epidermal or basic fibroblastic growth factor. (I) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a normal
 CC cell, an increase in the level of Cyr61 compared to the level in normal
 CC tissue indicates the presence of breast cancer. The level of Cyr61 being
 CC determined by exposing the tissues to (I), and an increase in the level
 CC of bound antibody by the suspect/positive cell as compared to the normal
 CC tissue indicates the presence of breast cancer. The present sequence
 CC represents the human Cyr61 protein, which is used in the exemplification
 CC of the present invention
 XX SQ Sequence 381 AA;
 PS Query Match 99.5%; Score 2106; DB 5; Length 381;
 CC Best Local Similarity 99.5%; Pred. No. 5.6e-161;

Qy	Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1	MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
1	MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
61	NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCVNSRIYQNGESFQPNCKHQ 120
61	NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCVNSRIYQNGESFQPNCKHQ 120
121	CTCIDGAVGCIPLCPQBELSLNLCNPNRLVKVTGQCCSEWVCDSEDSIKDPMEDQDGLLG 180
121	CTCIDGAVGCIPLCPQBELSLNLCNPNRLVKVTGQCCSEWVCDSEDSIKDPMEDQDGLLG 180
181	KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIQVTTWSQCS 240
181	KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIQVTTWSQCS 240
241	KTCGTGISTRVNDNPECLRVKTRICVVRPCGQPVYSSLLKGGKCKSKTKKSPPEVFRFTY 300
241	KTCGTGISTRVNDNPECLRVKTRICVVRPCGQPVYSSLLKGGKCKSKTKKSPPEVFRFTY 300
301	AGCLSVKKYRKYCGSSVDGRCCTPQLTRTVKMRFRCEGDTFSKNVMMIOSCKNCP 360
301	AGCLSVKKYRKYCGSSVDGRCCTPQLTRTVKMRFRCEGDTFSKNVMMIOSCKNCP 360
361	HANEAAPFFYRLFNDDIHKFRD 381
361	HANEAAPFFYRLFNDDIHKFRD 381

RESULT 7

AAE18107

ID AAE18107 standard; protein; 381 AA.

AAE18107;

AC

07-MAY-2002 (first entry)

Human connective tissue growth factor-2 (CTGF-2).

Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour; ischaemia; restenosis; tissue repair; wound healing; congenital defect; cardiovascular disease; atherosclerosis; heart failure; angina; trauma; burns; osteoporosis; periodontal disease; liver failure; tranquilizer; vulnary; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer; gene therapy.

Homo sapiens.

Key

Location/Qualifiers

35..41

/note= "Immunogenic epitope"

47..51

/note= "Immunogenic epitope"

59..75

/note= "Immunogenic epitope"

91..119

/note= "Immunogenic epitope"

145..150

/note= "Immunogenic epitope"

164..176

/note= "Immunogenic epitope"

202..208

/note= "Immunogenic epitope"

223..228

/note= "Immunogenic epitope"

239..244

/note= "Immunogenic epitope"

250..257

/note= "Immunogenic epitope"

279..296

/note= "Immunogenic epitope"

FT

FT

FT

FT

FT Region 307..314

/note= "Immunogenic epitope"

FT Region 318..323

/note= "Immunogenic epitope"

FT Region 337..343

/note= "Immunogenic epitope"

XX WO200204480-A2.

PN 17-JAN-2002.

XX 11-JUL-2001; 2001WO-US021799.

XX 11-JUL-2000; 2000US-0217402P.

XX 18-MAY-2001; 2001US-0291642P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (TRGE) TRANSGENE SA.

XX Li H, Adams MD, Calenda V, Fataccioli V;

XX WPI; 2002-171698/22.

XX N-PSDB; AAD29095.

DR Stimulating angiogenesis in a mammal preferably human having ischemia or

XX restenosis or is treated for limb revascularization, by administering

XX connective tissue growth factor-2 polypeptide or polynucleotide.

XX Example 1; Fig 1; 131pp; English.

XX The present invention relates to a method for stimulating angiogenesis in

XX a mammal. The method comprises administering a polynucleotide encoding

XX connective tissue growth factor-2 (CTGF-2) or an active fragment or its

XX derivative. The method is useful for stimulating angiogenesis in a mammal

XX preferably human having ischemia or restenosis or is treated for limb

XX revascularisation which is leg or arm. The invention is useful for

XX inhibiting tumour growth, where angiogenesis is utilised for enhancing

XX the repair of connective and support tissue, promoting the attachment,

XX fixation and stabilisation of tissue implants and enhancing wound

XX healing, hence is useful for treating cardiovascular disease e.g.

XX atherosclerosis, reperfusion injury such as heart failure, angina,

XX ischaemia; and is also used to differentiate, proliferate and attract

XX cells leading to regeneration of tissues which is utilised to repair

XX replace or protect tissue damaged by congenital defects, trauma (burns,

XX ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease, liver

XX failure), surgery including cosmetic plastic surgery. The present

XX sequence is human CTGF-2. CTGF-2 gene is useful in gene therapy

XX SQ

Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 5; Length 381;

Best Local Similarity 99.5%; Pred. No. 5.6e-161;

Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

Db 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCVNSRIYQNGESFQPNCKHQ 120

Db 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCVNSRIYQNGESFQPNCKHQ 120

Qy 121 CTCIDGAVGCIPLCPQBELSLNLCNPNRLVKVTGQCCSEWVCDSEDSIKDPMEDQDGLLG 180

Db 121 CTCIDGAVGCIPLCPQBELSLNLCNPNRLVKVTGQCCSEWVCDSEDSIKDPMEDQDGLLG 180

Qy 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIQVTTWSQCS 240

Db 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIQVTTWSQCS 240

Qy 241 KTCGTGISTRVNDNPECLRVKTRICVVRPCGQPVYSSLLKGGKCKSKTKKSPPEVFRFTY 300

Db 241 KTCGTGISTRVNDNPECLRVKTRICVVRPCGQPVYSSLLKGGKCKSKTKKSPPEVFRFTY 300

QY 301 AGCLSVKYPKYPKYGSCVDGRCTPQTRTVKRRFCEDGETFSKNVMMIOCKCNVNC 360
 DB 301 AGCLSVKYPKYPKYGSCVDGRCTPQTRTVKRRFCEDGETFSKNVMMIOCKCNVNC 360
 QY 361 HANEAAFPFYRLFNDFHIFKFRD 381
 DB 361 HANEAAFPFYRLFNDFHIFKFRD 381

RESULT 8
 ABU63222
 ID ABU63222 standard; protein; 381 AA.
 XX
 AC ABU63222;
 XX
 DT 17-OCT-2003 (first entry)
 XX
 DE Human connective tissue growth factor 2 (CTGF-2) #1.

XX Human; connective tissue growth factor 2; CTGF-2; support tissue;
 KW tissue repair; skin disorder; injury; acne; UV damage; burn;
 KW wrinkled skin; tissue implant; vulnerary; antiaging; dermatological;
 KW antiseborrheic.

XX Homo sapiens.
 OS
 PN US6534630-B1.
 XX
 PD 18-MAR-2003.

XX
 PF 08-JUL-1999; 99US-00348815.
 XX
 PR 12-JUL-1994; 94WO-US007736.
 PR 02-JUN-1995; 95US-00459101.

XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Li H, Adams MD;
 XX
 DR WPI; 2003-553676/52.
 DR N-PSDB; ACD26170.

XX New human connective tissue growth factor 2 polypeptides and
 PT polynucleotides useful for enhancing repair of connective and support
 PT tissues, and for treating skin disorders, e.g. injuries, acne, UV damage
 PT or burns, and wrinkled skin.
 XX
 PS Claim 1; Col 23-24; 20pp; English.

XX The present invention relates to the isolation of human connective tissue
 CC growth factor 2 (CTGF-2), and the polynucleotide sequence encoding it.
 CC The CTGF-2 polypeptide is useful for enhancing the repair of connective
 CC and support tissues, for treating skin disorders such as injuries, acne,
 CC UV damage or burns, and wrinkled skin, for promoting attachment, fixation
 CC and stabilisation of tissue implants, and as an immunogen to produce
 CC antibodies. The polynucleotide sequence encoding CTGF-2 may be used as a
 CC hybridisation probe for a cDNA library to isolate the full-length cDNA
 CC and other cDNAs, which have high sequence similarity to the CTGF-2 gene
 CC or similar biological activity. The polynucleotides are also useful for
 CC producing polypeptides by recombinant techniques, and in chromosome
 CC identification. The present sequence represents human CTGF-2. Note: The
 CC present sequence given as SEQ ID No:2 in the sequence listing differs
 CC from that given in Fig 1 (ABU63223)

XX Sequence 381 AA;
 SQ
 Query Match 99.5%; Score 2106; DB 6; Length 381;
 Best Local Similarity 99.5%; Pred. No. 5.6e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTLALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCKVCVKQL 60

DB 1 MSSRIARALAVVTLHLTLALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCKVCVKQL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRCEYNSRIYQNGESFQPNCHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRCEYNSRIYQNGESFQPNCHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRVLKVTGQCCBEWVDCDESIKDPMEDQDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRVLKVTGQCCBEWVDCDESIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTRNNELIATVAGKGRSLKELPVFGMEPRILYNPLOGQKCIQVOTTSWQCS 240
 DB 181 KELGFDASEVELTRNNELIATVAGKGRSLKELPVFGMEPRILYNPLOGQKCIQVOTTSWQCS 240
 QY 241 KTCGTGISTRTVNDNPECLRVKETRICEVRPCQGPVYSSLKGGKCKTKKSPPEPRTY 300
 DB 241 KTCGTGISTRTVNDNPECLRVKETRICEVRPCQGPVYSSLKGGKCKTKKSPPEPRTY 300
 QY 301 AGCLSVKYPKYPKYGSCVDGRCTPQTRTVKRRFCEDGETFSKNVMMIOCKCNVNC 360
 DB 301 AGCLSVKYPKYPKYGSCVDGRCTPQTRTVKRRFCEDGETFSKNVMMIOCKCNVNC 360
 QY 361 HANEAAFPFYRLFNDFHIFKFRD 381
 DB 361 HANEAAFPFYRLFNDFHIFKFRD 381

RESULT 9
 AAB43987
 ID AAB43987 standard; protein; 455 AA.
 XX
 AC AAB43987;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1432.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
 KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocrotropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.
 OS
 PN WO20005350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005882.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR N-PSDB; AAC78196.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 11; Page 2116-2118; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44339. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities

CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antiaerthmatic; antirheumatic; antiarthritic;
 CC antineoplastic; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells; to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX Sequence 455 AA;
 SQ

Query Match 99.5%; Score 2106; DB 3; Length 455;
 Best Local Similarity 99.5%; Pred. No. 7e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLAEPKAPGVGLVRDGGCCCKVCAKQL 60
 DB 75 MSSRIARALALVVTLLHLTRALSTCPAACHCPLAEPKAPGVGLVRDGGCCCKVCAKQL 134

QY 61 NEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCYNSRIYQNGSFOPNCOHQ 120
 DB 135 NEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCYNSRIYQNGSFOPNCKHQ 194

QY 121 CTCIDGAVGCIPLCQELSLNGLGCPNRLVKVTGQCCEWVCEDSIKDPMEDQDGLLG 180
 DB 195 CTCIDGAVGCIPLCQELSLNGLGCPNRLVKVTGQCCEWVCEDSIKDPMEDQDGLLG 254

QY 181 KELGFDASEVELTRNELIAVGKSLKRLPVFGMEPRILNPLQGGKCIQVTTWSQCS 240
 DB 255 KELGFDASEVELTRNELIAVGKSLKRLPVFGMEPRILNPLQGGKCIQVTTWSQCS 314

QY 241 KTCGTGISTRTVNDPECLVKETRICVRPGQPVYSSLLKKGKCKTKKSPPEVFRTY 300
 DB 315 KTCGTGISTRTVNDPECLVKETRICVRPGQPVYSSLLKKGKCKTKKSPPEVFRTY 374

QY 301 AGCLSVKXKYPKYGCSVDGRCCTPQLTRTKMFRCEGDTFSKNVMIOSCKYVNCNP 360
 DB 375 AGCLSVKXKYPKYGCSVDGRCCTPQLTRTKMFRCEGDTFSKNVMIOSCKYVNCNP 434

QY 361 HANEAAFFPYRLFNDIHKFRD 381
 DB 435 HANEAAFFPYRLFNDIHKFRD 455

RESULT 10
 ABG76937
 ID ABG76937 standard; protein; 381 AA.
 XX
 AC ABG76937;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human protein, comprising CYR61, designated SEC1.
 XX
 KW Human; SEC; NOV; immunosuppressive; hepatotropic; antiinflammatory;
 KW angiotensin-associated disorder; diagnostic; gene therapy;
 KW developmental disorder; immune disease;
 KW signal transduction pathway disorder; metabolic disorder;
 KW feeding disorder; obesity; wasting disorder; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;
 KW asthma; atherosclerosis; cardiomyopathy; angina pectoris;
 KW autoimmune disease; retinal disease; cirrhosis; diabetes;
 KW infectious disease; human immunodeficiency virus; HIV; cancer;

KW hypertension; hypotension; multiple sclerosis; urinary retention;
 KW osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;
 KW haemophilia; cirrhosis; immunogen; vaccine.
 OS Homo sapiens.
 XX WO200255705-A2.
 PN 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-US000609.
 XX 11-JAN-2001; 2001US-0261013P.
 PR 11-JAN-2001; 2001US-0261014P.
 PR 11-JAN-2001; 2001US-0261018P.
 PR 11-JAN-2001; 2001US-0261026P.
 PR 11-JAN-2001; 2001US-0261029P.
 PR 17-AUG-2001; 2001US-0313170P.
 PR 10-SEP-2001; 2001US-0318410P.
 XX (CURA-) CURAGEN CORP.
 XX Mezes PS, Rastelli L, Herrmann JL, Macdougall JR, Zhong H;
 XX Casman SJ, Beldog F, Shinkets RA, Gorman L, Craata OR, Mysore KK;
 PI Folkerts O, Martin GB, Eisen A, Spaderna SK, Vernet CAM, Bergh C;
 PI Spytek RA, Dipippo VA, Zernusen BD, Peyman JA, Ellerman K, Stone DJ;
 PI Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
 PI Edinger S;
 XX WPI; 2002-590675/63.
 DR N-PSDB; AB959522.
 XX Human SECX/NOVX polypeptide useful for diagnosing, preventing or treating
 PT disorders associated with aberrant expression or activity of SECX/NOVX
 PT nucleic acids and proteins e.g., diabetes.
 XX Example 3; Page 9; 443pp; English.
 XX The invention discloses the isolated human polypeptides, and
 CC polynucleotides encoding them, that have been designated SECX and NOVX.
 CC The polypeptides can be used for treating, or delaying, the onset of an
 CC angiotensin-associated disorder or treating a pathological state in a
 CC subject, preferably a mammal. They can also be used in determining the
 CC presence of, or predisposition to, a disease associated with altered
 CC levels of the polypeptides and polynucleotides of any one of the 12
 CC sequences (SEC1-12), for raising antibodies for identifying an agent
 CC that binds to, or that modulates the expression or activity of the
 CC polypeptide, for treating or preventing a NOVX-associated disorder (NOV1-
 CC 8) and as a pharmaceutical composition comprising the polypeptide,
 CC polynucleotide or the antibody. The polypeptides and polynucleotides are
 CC useful in diagnostic applications where their amounts are assessed, or
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or
 CC preventing disorders or syndromes such as developmental disorders, immune
 CC diseases, signal transduction pathway disorders, metabolic disorders,
 CC feeding disorders (including obesity), wasting disorders,
 CC neurodegenerative disorders (including Alzheimer's disease and
 CC Parkinson's disease), behavioural disorders, allergies, asthma,
 CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
 CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
 CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
 CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
 CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
 CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They may
 CC also be used as immunogens to produce antibodies specific for the
 CC invention, and as vaccines. Further, they are useful for screening
 CC potential agonist and antagonist compounds. The sequences presented in
 CC ABG76937-ABG76956 are the human SEC1-12 and NOV1-8 proteins
 XX Sequence 381 AA;
 SQ

Query Match 99.1%; Score 2098; DB 5; Length 381;
 Best Local Similarity 99.0%; Pred. No. 2.5e-160;
 Matches 377; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLRLALSTCPAACHCPLAPKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALALVVTLLHLRLALSTCPAACHCPLAPKCAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGOCCEWVCDSDSIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGOCCEWVCDSDSIKDPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIAVKGSLKRLPVFGMEPRILYNPLQOGKCIQVOTTSWQS 240
 DB 181 KELGFDASEVELTRNNELIAVKGSLKRLPVFGMEPRILYNPLQOGKCIQVOTTSWQS 240
 QY 241 KTCGTGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFTY 300
 DB 241 KTCGTGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFTY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPOLTTRTVKMFRCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPKYCGSCVDGRCTPOLTTRTVKMFRCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAAFPFVRLFNDFHFRD 381
 DB 361 HANEAAFPFVRLFNDFHFRD 381

RESULT 11

ADB75260
 ID ADB75260 standard; protein; 381 AA.

AC ADB75260;

XX 04-DEC-2003 (first entry)

DT Prostate cancer marker protein.

DE Prostate; cancer; cytostatic; gene therapy; marker.

XX Homo sapiens.

XX WO2003009814-A2.

XX 06-FEB-2003.

XX 25-JUL-2002; 2002WO-US023913.

XX 25-JUL-2001; 2001US-0307982P.

XX 22-AUG-2001; 2001US-0314356P.

XX 25-SEP-2001; 2001US-0325020P.

XX 12-DEC-2001; 2001US-0341746P.

XX 05-MAR-2002; 2002US-0362158P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

XX Hoersch S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;

XX WPI; 2003-248033/24.

XX New nucleic acid molecule, useful for diagnosing or treating prostate

XX cancer.

XX Disclosure; SEQ ID NO 84; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with

XX the cancerous state of prostate cells. Also disclosed is a method of

XX assessing whether a patient is afflicted with prostate cancer. The method

XX of the invention involves assessing whether a patient is afflicted with

XX prostate cancer by comparing the level of expression of a marker in a

CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 381 AA;

Query Match 99.1%; Score 2098; DB 7; Length 381;

Best Local Similarity 99.0%; Pred. No. 2.5e-160;

Matches 377; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLRLALSTCPAACHCPLAPKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALALVVTLLHLRLALSTCPAACHCPLAPKCAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFQPNCHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGOCCEWVCDSDSIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGOCCEWVCDSDSIKDPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIAVKGSLKRLPVFGMEPRILYNPLQOGKCIQVOTTSWQS 240
 DB 181 KELGFDASEVELTRNNELIAVKGSLKRLPVFGMEPRILYNPLQOGKCIQVOTTSWQS 240
 QY 241 KTCGTGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFTY 300
 DB 241 KTCGTGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFTY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPOLTTRTVKMFRCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPKYCGSCVDGRCTPOLTTRTVKMFRCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAAFPFVRLFNDFHFRD 381
 DB 361 HANEAAFPFVRLFNDFHFRD 381

RESULT 12

ABB09202
 ID ABB09202 standard; protein; 374 AA.

AC ABB09202;

XX 08-JUL-2002 (first entry)

DT HCGF CNN family protein sequence SEQ ID NO:12.

DE Human, small CCN-like growth factor; SCGF; vulnary; osteopathic;

XX Gene therapy; muscle wasting disease; osteoporosis; wound healing;

XX tissue regeneration; angiogenesis.

XX Unidentified.

XX US2002049304-A1.

XX 25-APR-2002.

XX 14-MAY-2001; 2001US-00853625.

XX 06-JUN-1995; 95US-00468847.

XX 01-APR-1998; 98US-00053587.

XX (HAST/) HASTINGS G A.

XX (ADAM/) ADAMS M D.

XX Hastings GA, Adams MD;
 XX WPI; 2002-382150/41.
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 XX growth factor, useful for treating muscle wasting disease, and
 XX osteoporosis.
 XX Disclosure; Fig 2A-D; 33pp; English.
 XX The present invention describes human small CCN-like growth factor
 XX (SCGF). SCGF has vulnary and osteopathic activities, and can be used in
 XX gene therapy. The SCGF polypeptides and polynucleotides can be used for
 XX treating muscle wasting diseases and osteoporosis, and to stimulate
 XX wound healing and tissue regeneration, to promote angiogenesis and to
 XX stimulate proliferation of vascular, smooth muscle and endothelial cell
 XX production. The present sequence represents a CCN family protein which is
 XX given in comparison with the human SCGF in the exemplification of the
 XX present invention
 XX Sequence 374 AA;
 SQ
 Query Match 93.2%; Score 1971.5; DB 5; Length 374;
 Best Local Similarity 95.7%; Pred. No. 3.6e-150;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCKTOPCDHTKGLNFCGASSTALKGICRAQSEGRPCVYNSRIYONGESFQNCQHQ 120
 Db 61 NEDCKTOPCDHTKGLNFCGASSTALKGICRAQSEGRPCVYNSRIYONGESFQNCQHQ 120
 QY 121 CTCIDGAVG-CIPLCPQELSLPNCNPNRLVKVTGQCCEWVCDSDSIKDPMEDQDGLL 179
 Db 121 CTCIDGAVG-CIPLCPQELSLPNCNPNRLVKVTGQCCEWVCDSDSIKDPMEDQDGLL 180
 QY 180 GKEIGFDASEVELTNNELIANGKRSIKRLPVFGMEPRILYNPLOGKCIQVOTTSWQC 239
 Db 181 GKGLGFDASEVELTNNELIANGKRSIKRLPVFGMEPRILYNPLOGKCIQVOTTSWQC 240
 QY 240 SKTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKSKTKKSPPEVRFT 299
 Db 241 SKTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKSKTKKSPPEVRFT 300
 QY 300 YAGCLSVKXKRPKYCGSCVDGRCCTPQLTRVTKMRFCEDGETFSKNVMMIQSKCNVNC 359
 Db 301 YAGCLSVKXKRPKYCGSCVDGRCCTPQLTRVTKMRFCEDGETFSKNVMMIQSKCNVNC 360
 QY 360 PHANEAAFPFVRLF 373
 Db 361 PHANEAAFPFVRLF 374
 RESULT 13
 AAR90919
 ID AAR90919 standard; protein; 375 AA.
 XX
 AC AAR90919;
 XX
 DT 25-MAR-2003 (revised)
 DT 25-JUN-1996 (first entry)
 XX
 DE Connective tissue growth factor-2.
 XX CTGF-2; connective tissue growth factor-2; secreted protein;
 KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 KW insulin-like growth factor; fibroblast growth factor; Cry61.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide
 FT Protein 25..375
 FT /label= mature_protein
 XX WO9601896-A1.
 PN 25-JAN-1996.
 XX 12-JUL-1994; 94WO-US007736.
 PF 12-JUL-1994; 94WO-US007736.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Li H, Adams MD;
 PI WPI; 1996-097626/10.
 DR N-PSDB; AAT12653.
 XX Connective tissue growth factor-2 and DNA encoding it - useful to enhance
 PT the repair of connective and support tissue, and to enhance wound
 PT healing.
 XX Claim 1; Fig 1A-C; 46pp; English.
 PS
 CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family of
 CC growth factors which include IGF (insulin-like growth factor), PDGF
 CC (platelet-derived growth factor), and FGF (fibroblast growth factor).
 CC CTGF-2 exhibits 89 percent identity and 93 percent similarity to Cry61.
 CC Cry61 is a growth factor-inducible immediate early gene initially
 CC identified in serum-stimulated mouse fibroblasts. It encodes a member of
 CC an emerging family of secreted proteins which are also a group of
 CC cysteine-rich proteins. This group of GPs are important for normal
 CC growth, differentiation, morphogenesis of the cartilaginous skeleton of
 CC an embryo and cell growth. (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 375 AA;
 SQ
 Query Match 93.2%; Score 1971.5; DB 2; Length 375;
 Best Local Similarity 95.7%; Pred. No. 3.6e-150;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCKTOPCDHTKGLNFCGASSTALKGICRAQSEGRPCVYNSRIYONGESFQNCQHQ 120
 Db 61 NEDCKTOPCDHTKGLNFCGASSTALKGICRAQSEGRPCVYNSRIYONGESFQNCQHQ 120
 QY 121 CTCIDGAVG-CIPLCPQELSLPNCNPNRLVKVTGQCCEWVCDSDSIKDPMEDQDGLL 179
 Db 121 CTCIDGAVG-CIPLCPQELSLPNCNPNRLVKVTGQCCEWVCDSDSIKDPMEDQDGLL 180
 QY 180 GKEIGFDASEVELTNNELIANGKRSIKRLPVFGMEPRILYNPLOGKCIQVOTTSWQC 239
 Db 181 GKGLGFDASEVELTNNELIANGKRSIKRLPVFGMEPRILYNPLOGKCIQVOTTSWQC 240
 QY 240 SKTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKSKTKKSPPEVRFT 299
 Db 241 SKTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKSKTKKSPPEVRFT 300
 QY 300 YAGCLSVKXKRPKYCGSCVDGRCCTPQLTRVTKMRFCEDGETFSKNVMMIQSKCNVNC 359
 Db 301 YAGCLSVKXKRPKYCGSCVDGRCCTPQLTRVTKMRFCEDGETFSKNVMMIQSKCNVNC 360
 QY 360 PHANEAAFPFVRLF 373
 Db 361 PHANEAAFPFVRLF 374

RESULT 14
 AAY31620
 ID AAY31620 standard; protein; 375 AA.
 XX AAY31620;
 AC AAY31620;
 DT 02-NOV-1999 (first entry)
 XX Human CTGF-2.
 DE
 KW Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
 KW skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= Signal_peptide
 FT Protein 25..375
 FT Misc-difference 268
 FT /note= "Cys encoded by ICT"
 XX
 PN US5945300-A.
 XX
 PD 31-AUG-1999.
 XX
 PF 02-JUN-1995; 95US-00459101.
 XX
 PR 12-JUL-1994; 94WO-US007736.
 XX
 PA (ADAM/) ADAMS M D.
 PA (LIHH/) LI H.
 XX
 PI Adams MD, Li H;
 XX
 DR WPI; 1999-508171/42.
 DR N-PSDB; AA211720.
 XX
 PT Polynucleotides encoding growth factor polypeptides useful for enhancing
 PT the repair of connective tissue and support tissue.
 XX
 PS Claim 1; Fig 1; 20pp; English.
 CC This sequence represents human connective tissue growth factor-2 (CTGF-
 CC 2). CTGF-2 cDNA was isolated from a cDNA library derived from human
 CC foetal lung. In one instance, the cDNA was cloned into a baculovirus
 CC expression vector, having first been amplified and modified via PCR using
 CC primers AA211721 and AA211722. In another instance, the cDNA was cloned
 CC into a COS cell expression vector, with prior amplification and
 CC modification using PCR primers AA211723 and AA211724. CTGF-2 is
 CC structurally and functionally related to a family of growth factors which
 CC include IGF (insulin-like growth factor), PDGF (platelet-derived growth
 CC factor) and FGF (fibroblast growth factor). This emerging family of
 CC cysteine-rich secreted proteins are important for normal growth,
 CC differentiation, morphogenesis of the cartilaginous skeleton of an embryo
 CC and cell growth. Their functions also include wound healing, tissue
 CC repair, implant fixation and stimulating increased bone mass. CTGF-2 may
 CC be used to enhance the repair of connective tissue and support tissue and
 CC can therefore treat skin disorders e.g., acne, aging, UV damage or burns.
 CC CTGF-2 can be used to promote the attachment, fixation and stabilization
 CC of tissue implants inserted during reconstructive surgery, and can be
 CC used to enhance the healing of external wounds. It can be used in the
 CC treatment of injured or depleted bone as it promotes the growth of
 CC connective tissue, bone and cementum and stimulates protein and collagen
 CC synthesis
 XX
 SQ Sequence 375 AA;
 Query Match 93.2%; Score 1971.5; DB 2; Length 375;
 Best Local Similarity 95.7%; Pred. No. 3.6e-150;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 PS Disclosure; Fig 11; 131pp; English.
 CC The present invention relates to a method for stimulating angiogenesis in

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLKAPKAPGVGLVRDGGCGCKVCAKOL 60
 DB 1 MSSRIVRELAVVTLHLTRVGLSTCPADCHCPLKAPKAPGVGLVRDGGCGCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRCEYNSRYQNGESFQPNCOHQ 120
 DB 61 NEDCRKTQPCDHTKGLCNFGASSTALXGICRAQSEGRCEYNSRYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCPQELSLNLCNPNRLVKVTGQCCCEWVCDSDSIKDPMEDQDGLL 179
 DB 121 CTCIGWRRGACIPLCPQELSLNLCNPNRLVKVTGQCCCEWVCDSDSIKDPMEDQDGLL 180
 QY 180 GKELGFDASEVELTRNNELIANGKESLXBLVFGNEPRILYNPLOGKQKCIQVTTWSOC 239
 DB 181 GKGLGFDASEVELTRNNELIANGKESLXBLVFGNEPRILYNPLOGKQKCIQVTTWSOC 240
 QY 240 SKTCGTGIGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKKKKCSKTKSPPEVRF 299
 DB 241 SKTCGTGIGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKKKKCSKTKSPPEVRF 300
 QY 300 YAGCLSVKKYRKYCGSCVDGRCCCTPOLTRTVKMFRCEDGETFSKNVMMIOSCKNYNC 359
 DB 301 YAGCLSVKKYRKYCGSCVDGRCCCTPOLTRTVKMFRCEDGETFSKNVMMIOSCKNYNC 360
 QY 360 PHANEAAPFFYRLF 373
 DB 361 PHANEAAPFFYRLF 374
 RESULT 15
 AA218108
 ID AA218108 standard; protein; 375 AA.
 XX AA218108;
 XX 07-MAY-2002 (first entry)
 DE Human alternative connective tissue growth factor-2 (CTGF-2).
 XX Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
 KW ischaemia; restenosis; tissue repair; wound healing; congenital defect;
 KW cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
 KW burns; osteoporosis; periodontal disease; liver failure; tranquillizer;
 KW vulnary; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
 KW gene therapy.
 XX Homo sapiens.
 OS WO200204480-A2.
 XX 17-JAN-2002.
 XX 11-JUL-2001; 2001WO-US021799.
 PR 11-JUL-2000; 2000US-0217402P.
 PR 18-MAY-2001; 2001US-0291642P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (TRGE) TRANSGENE SA.
 XX Li H, Adams MD, Calenda V, Fataccioli V;
 DR WPI; 2002-171698/22.
 DR N-PSDB; AAD29039.
 XX Stimulating angiogenesis in a mammal preferably human having ischemia or
 PT restenosis or is treated for limb revascularization, by administering
 PT connective tissue growth factor-2 polypeptide or polynucleotide.
 XX
 PS Disclosure; Fig 11; 131pp; English.
 CC The present invention relates to a method for stimulating angiogenesis in

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:54:44 ; Search time 16.5434 Seconds
(without alignments)
1188.962 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALAVVTLHLTR.....ANEAPFFVRLFNHDKRD 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	381	4	US-09-142-569-4
2	2106	99.5	381	4	US-09-348-815-2
3	1971.5	93.2	374	1	US-08-468-847B-12
4	1964.5	92.8	375	2	US-08-459-101A-2
5	1929	91.2	373	1	US-08-468-847B-11
6	1929	91.2	373	4	US-09-142-569-2
7	1695	80.1	375	1	US-08-468-847B-13
8	953	45.0	347	4	US-09-582-337-2
9	952.5	45.0	348	4	US-09-292-036-3
10	948.5	44.8	348	1	US-08-468-847B-15
11	948.5	44.8	347	4	US-09-142-569-6
12	948	44.8	347	4	US-09-187-478-2
13	948	44.8	347	4	US-09-232-036-2
14	947	44.8	349	1	US-08-167-828-2
15	947	44.8	349	1	US-08-386-680-2
16	947	44.8	349	1	US-08-459-717-2
17	947	44.8	349	1	US-08-712-302-2
18	947	44.8	349	2	US-08-880-031-2
19	947	44.8	349	3	US-09-054-368-2
20	947	44.8	349	3	US-09-037-179-2
21	947	44.8	349	3	US-09-034-274-2
22	947	44.8	349	3	US-09-080-715-2
23	947	44.8	349	3	US-09-056-704-2
24	947	44.8	349	4	US-09-292-036-4
25	947	44.8	349	4	US-09-253-316-26
26	947	44.8	349	4	US-09-142-569-8
27	947	44.8	349	4	US-09-461-688-2

28 947 44.8 349 5 PCT-US96-08140-2 Sequence 2, Appli
29 938.5 44.4 348 1 US-08-468-847B-14 Sequence 14, Appli
30 851.5 40.2 351 1 US-08-468-847B-16 Sequence 16, Appli
31 827.5 39.1 357 1 US-08-468-847B-17 Sequence 17, Appli
32 827.5 39.1 357 4 US-09-253-316-25 Sequence 25, Appli
33 767 36.2 367 4 US-09-182-145-4 Sequence 4, Appli
34 767 36.2 367 4 US-09-182-145-8 Sequence 8, Appli
35 766 36.2 367 4 US-09-182-145-7 Sequence 7, Appli
36 766 36.2 367 4 US-09-182-145-22 Sequence 22, Appli
37 760.5 35.9 345 4 US-09-182-145-3 Sequence 3, Appli
38 760.5 35.9 345 4 US-09-182-145-6 Sequence 6, Appli
39 759.5 35.9 345 4 US-09-182-145-5 Sequence 5, Appli
40 759.5 35.9 345 4 US-09-182-145-21 Sequence 21, Appli
41 758 35.8 367 4 US-09-182-145-11 Sequence 12, Appli
42 754.5 35.7 345 4 US-09-182-145-11 Sequence 11, Appli
43 613 29.0 339 4 US-09-182-145-36 Sequence 36, Appli
44 613 29.0 354 4 US-09-182-145-37 Sequence 37, Appli
45 613 29.0 354 4 US-09-253-316-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-142-569-4
; Sequence 4, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/142,569
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: "Human Cyt-61 amino acid sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-142-569-4

Query Match 100.0%; Score 2116; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 8.1e-177;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSSRIARALAVVTLHLTRALSTCPAACHPCAPGVLRDCCCKVCAKQL 60
QY |||||

Db 1 MSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCVNSRIYONGESFQPNCOHQ 120
 Db 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCVNSRIYONGESFQPNCOHQ 120
 QY 121 CTCIDGAVGCIPLCQELSLPGLGCPNRLVKTQCCCEEWVCDSDSIKDPMEDQDGLLG 180
 Db 121 CTCIDGAVGCIPLCQELSLPGLGCPNRLVKTQCCCEEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGPDASEVELTRNNELIANGKSLKELPVFGMEPRILYNPLOGQKCIQVTTSSQCS 240
 Db 181 KELGPDASEVELTRNNELIANGKSLKELPVFGMEPRILYNPLOGQKCIQVTTSSQCS 240
 QY 241 KTCGTGISTRVNDNPECLVKETRICVPRGCGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 Db 241 KTCGTGISTRVNDNPECLVKETRICVPRGCGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPQTRTVORFRCEDETFSSKNVMMIOSCKNYNCP 360
 Db 301 AGCLSVKKYRPKYCGSCVDGRCTPQTRTVORFRCEDETFSSKNVMMIOSCKNYNCP 360
 QY 361 HANEAAPFPYRLFNDIHKFRD 381
 Db 361 HANEAAPFPYRLFNDIHKFRD 381

RESULT 2

US-09-348-815-2
 ; Sequence 2, Application US/09348815
 ; Patent No. 6534630
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, HAODONG
 ; ADAMS, MARK D
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
 ; STREET: 9410 KEY WEST AVENUE
 ; CITY: ROCKVILLE
 ; STATE: MD
 ; COUNTRY: US
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/348,815
 ; FILING DATE: 08-Jul-1999
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: JONATHAN L. KLEIN
 ; REGISTRATION NUMBER: 41,119
 ; REFERENCE/DOCKET NUMBER: PF126PID1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 301-309-8504
 ; TELEFAX: 301-309-8439
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 381 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-348-815-2

Query Match 99.5%; Score 2106; DB 4; Length 381;
 Best Local Similarity 99.5%; Pred. No. 66-176;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKQL 60

Db 1 MSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCVNSRIYONGESFQPNCOHQ 120
 Db 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCQELSLPGLGCPNRLVKTQCCCEEWVCDSDSIKDPMEDQDGLLG 180
 Db 121 CTCIDGAVGCIPLCQELSLPGLGCPNRLVKTQCCCEEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGPDASEVELTRNNELIANGKSLKELPVFGMEPRILYNPLOGQKCIQVTTSSQCS 240
 Db 181 KELGPDASEVELTRNNELIANGKSLKELPVFGMEPRILYNPLOGQKCIQVTTSSQCS 240
 QY 241 KTCGTGISTRVNDNPECLVKETRICVPRGCGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 Db 241 KTCGTGISTRVNDNPECLVKETRICVPRGCGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 QY 301 AGCLSVKKYRPKYCGSCVDGRCTPQTRTVORFRCEDETFSSKNVMMIOSCKNYNCP 360
 Db 301 AGCLSVKKYRPKYCGSCVDGRCTPQTRTVORFRCEDETFSSKNVMMIOSCKNYNCP 360
 QY 361 HANEAAPFPYRLFNDIHKFRD 381
 Db 361 HANEAAPFPYRLFNDIHKFRD 381

RESULT 3

US-08-468-847B-12
 ; Sequence 12, Application US/08468847B
 ; Patent No. 5780263
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,847B
 ; FILING DATE: 6 June 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MULLINS, J.G.
 ; REGISTRATION NUMBER: 33,073
 ; REFERENCE/DOCKET NUMBER: 325800-442
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 374 AMINO ACIDS
 ; TYPE: AMINO ACID
 ; STRANDEDNESS:
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: PROTEIN
 US-08-468-847B-12

Query Match 93.2%; Score 1971.5; DB 1; Length 374;

Best Local Similarity 95.7%; Pred. No. 3.3e-164;
Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
QY 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDCKTQPCDHTKGLCNFNGASSTALKGI CRAQSEGRPCYNRIYONGESFQPNCOHQ 120
DB 61 NEDCKTQPCDHTKGLCNFNGASSTALKGI CRAQSEGRPCYNRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVG-CIPLCPQELSLPGLCPNRLVKVTGQCCCEWVDEDSIKDPMEDQDGLL 179
DB 121 CTCIDGAVG-CIPLCPQELSLPGLCPNRLVKVTGQCCCEWVDEDSIKDPMEDQDGLL 180
QY 180 GKELGFDASEVELTRNNELIANGKRSKRLPVFGMEPRILYNPLOGKCIQVTTWSQC 239
DB 181 GKELGFDASEVELTRNNELIANGKRSKRLPVFGMEPRILYNPLOGKCIQVTTWSQC 240
QY 240 SKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKKGKSKTKKSPVPVFT 299
DB 241 SKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKKGKSKTKKSPVPVFT 300
QY 300 YAGCLSVKKYRKYCGSCVDRCCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNINC 359
DB 301 YAGCLSVKKYRKYCGSCVDRCCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNINC 360
QY 360 PHANEAAPFFYRLF 373
DB 361 PHANEAAPFFYRLF 374

RESULT 4

US-08-459-101A-2
Sequence 2, Application US/08459101A
Patent No. 5945300
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Connective Tissue Growth Factor-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESS: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,101A
FILING DATE: June 2, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07736
FILING DATE: 12 JUL 94
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-317
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN
US-08-459-101A-2

Query Match 92.8%; Score 1964.5; DB 2; Length 375;
Best Local Similarity 95.5%; Pred. No. 1.3e-163;
Matches 357; Conservative 2; Mismatches 14; Indels 1; Gaps 1;
QY 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDCKTQPCDHTKGLCNFNGASSTALKGI CRAQSEGRPCYNRIYONGESFQPNCOHQ 120
DB 61 NEDCKTQPCDHTKGLCNFNGASSTALKGI CRAQSEGRPCYNRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVG-CIPLCPQELSLPGLCPNRLVKVTGQCCCEWVDEDSIKDPMEDQDGLL 179
DB 121 CTCIDGAVG-CIPLCPQELSLPGLCPNRLVKVTGQCCCEWVDEDSIKDPMEDQDGLL 180
QY 180 GKELGFDASEVELTRNNELIANGKRSKRLPVFGMEPRILYNPLOGKCIQVTTWSQC 239
DB 181 GKELGFDASEVELTRNNELIANGKRSKRLPVFGMEPRILYNPLOGKCIQVTTWSQC 240
QY 240 SKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKKGKSKTKKSPVPVFT 299
DB 241 SKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKKGKSKTKKSPVPVFT 300
QY 300 YAGCLSVKKYRKYCGSCVDRCCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNINC 359
DB 301 YAGCLSVKKYRKYCGSCVDRCCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNINC 360
QY 360 PHANEAAPFFYRLF 373
DB 361 PHANEAAPFFYRLF 374

RESULT 5

US-08-468-847B-11
Sequence 11, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESS: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

; LENGTH: 379 AMINO ACIDS
 ; TYPE: AMINO ACID
 ; STRANDEDNESS:
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: PROTEIN
 ; US-08-468-847B-11

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred. No. 1.7e-160;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY	1	MSRIARALAVVTLHLRLALSTCPAACHCPLKAPKAPGVGLVRDGGCGCKVCAKQL	60
DB	1	MSSTFRTILAVVTLHLRLALSTCPAACHCPLKAPKAPGVGLVRDGGCGCKVCAKQL	60
QY	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCVNSRIYQNGESFQPCNHQ	120
DB	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCVNSRIYQNGESFQPCNHQ	120
QY	121	CTCIDGAVGICPLCPQELSPLNLCNPNRLVKTGQCCEWVCDSDIKDPMEDDGLLG	180
DB	121	CTCIDGAVGICPLCPQELSPLNLCNPNRLVKTGQCCEWVCDSDIKDPMEDDGLLG	180
QY	181	KELGPDASEVELTRNNELIAVKGKSLKRLPVFGMEPRILYNPL--OGOKCIVQTTWSQ	238
DB	179	--LGLDASEVELTRNNELIAVKGKSLKRLPVFGTEPRVFLPFLHAHQKQKCIQVTTWSQ	236
QY	239	CSKTCGTGISTRVTNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKTKKSPVPVF	298
DB	237	CSKSCGTGISTRVTNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKTKKSPVPVF	296
QY	299	TYAGCLSVKKYRKYKCGSCVDGRCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNYN	358
DB	297	TYAGCSVKYRKYKCGSCVDGRCCTPQLTRTVKMFRCEDGEMFSKNVMIQSKCNYN	356
QY	359	CPHANEAAFPFYRLFNDIHKFRD 381	
DB	357	CPHPNEASFRLYSLFNDIHKFRD 379	

RESULT 6

US-09-142-569-2

; Sequence 2, Application US/09142569

; Patent No. 6413735

; GENERAL INFORMATION:

; APPLICANT: Lau, Lester F.

; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/142,569

; FILING DATE: 02-Apr-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Clough, David W.

; REGISTRATION NUMBER: 36,107

; REFERENCE/DOCKET NUMBER: 28758/33766

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: "Mouse Cvr61 amino acid sequence"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-142-569-2
 Query Match 91.2%; Score 1929; DB 4; Length 379;
 Best Local Similarity 90.9%; Pred. No. 1.7e-160;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY	1	MSRIARALAVVTLHLRLALSTCPAACHCPLKAPKAPGVGLVRDGGCGCKVCAKQL	60
DB	1	MSSTFRTILAVVTLHLRLALSTCPAACHCPLKAPKAPGVGLVRDGGCGCKVCAKQL	60
QY	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCVNSRIYQNGESFQPCNHQ	120
DB	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCVNSRIYQNGESFQPCNHQ	120
QY	121	CTCIDGAVGICPLCPQELSPLNLCNPNRLVKTGQCCEWVCDSDIKDPMEDDGLLG	180
DB	121	CTCIDGAVGICPLCPQELSPLNLCNPNRLVKTGQCCEWVCDSDIKDPMEDDGLLG	180
QY	181	KELGPDASEVELTRNNELIAVKGKSLKRLPVFGMEPRILYNPL--OGOKCIVQTTWSQ	238
DB	179	--LGLDASEVELTRNNELIAVKGKSLKRLPVFGTEPRVFLPFLHAHQKQKCIQVTTWSQ	236
QY	239	CSKTCGTGISTRVTNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKTKKSPVPVF	298
DB	237	CSKSCGTGISTRVTNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKTKKSPVPVF	296
QY	299	TYAGCLSVKKYRKYKCGSCVDGRCCTPQLTRTVKMFRCEDGETFSKNVMIQSKCNYN	358
DB	297	TYAGCSVKYRKYKCGSCVDGRCCTPQLTRTVKMFRCEDGEMFSKNVMIQSKCNYN	356
QY	359	CPHANEAAFPFYRLFNDIHKFRD 381	
DB	357	CPHPNEASFRLYSLFNDIHKFRD 379	

RESULT 7

US-08-468-847B-13

; Sequence 13, Application US/08468847B

; Patent No. 5780263

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

; TITLE OF INVENTION: Human CCN-Like Growth Factor

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468,847B

; FILING DATE: 6 June 1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

```
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-13

Query Match      80.1%; Score 1695; DB 1; Length 375;
Best Local Similarity 81.0%; Pred. No. 4.5e-140;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSSRIARALALVLTLLHLTRAL-STCPAAACHCPLEAPKCAPGVGLVRDGCCKVCAKQ 59
DB 1 MGSAGARP-ALAAALLCLARLALGSPCPAVCQCPAAAPQCAPGVGLVDPDGCCKVCAKQ 59
QY 60 LNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPNCOH 119
DB 60 LNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPNCKH 119
QY 120 QCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQDGLL 179
DB 120 QCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQDGLL 179
QY 180 GKELGPDASEVELTRNNELIANGKRSKLEL-PVFGMEPRILYNPLOGQKCIQVTTWS 237
DB 178 SKFGLDASEGELTRNNELIANGKRSKLEL-PVFGMEPRILYNPLOGQKCIQVTTWS 232
QY 238 QCSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 297
DB 233 QCSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 292
QY 298 FTYAGCLSVKKYRKYKCGSCVDGRCCCTPQTRTVKRFCEDETFSKNVMIIQSCCKNY 357
DB 293 FTYAGCLSVKKYRKYKCGSCVDGRCCCTPQTRTVKRFCEDETFSKNVMIIQSCCKNY 352
QY 358 NCPHANEAAFPFYFLFNDIHKFRD 381
DB 353 NCPHANE-AAFPFYFLVNDIHKFRD 375

RESULT 8
US-09-582-337-2
; Sequence 2, Application US/09582337
; Patent No. 6562618
; GENERAL INFORMATION:
; APPLICANT: Japan Tobacco, Inc.
; TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor
; FILE REFERENCE: JP-009PCT
; CURRENT APPLICATION NUMBER: US/09/582,337
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: JP P1997-367699
; PRIOR FILING DATE: 1997-12-25
; PRIOR APPLICATION NUMBER: JP P1998-356183
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Rat
US-09-582-337-2

Query Match      45.0%; Score 953; DB 4; Length 347;
Best Local Similarity 46.2%; Pred. No. 2.5e-75;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSSRIARALALVLTLLHL-TRIAL-STCPAAACHCPLE-APKCAPGVGLVRDGCCKVCA 57
DB 1 MLASVAGPISLALVLLALCTRPATQDCSAQCQCAEAAPHPAGVSLVLDGCGCCRVCA 60
QY 58 KOLNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPN 117
DB 61 KOLGELCTERDPCDPHKGFLCDFGSPANRKGIVGTAK-DGAPCVFGSGVYSGSFSOSSC 119
QY 118 QHQCCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQD 177
DB 120 KYQCTCLDGAAGCVPLCSMDVRLSPDCPFPRRVKLPKGCCEEWVCDSDSIKDPMEDQD 168

; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-13

Query Match      80.1%; Score 1695; DB 1; Length 375;
Best Local Similarity 81.0%; Pred. No. 4.5e-140;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSSRIARALALVLTLLHLTRAL-STCPAAACHCPLEAPKCAPGVGLVRDGCCKVCAKQ 59
DB 1 MGSAGARP-ALAAALLCLARLALGSPCPAVCQCPAAAPQCAPGVGLVDPDGCCKVCAKQ 59
QY 60 LNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPNCOH 119
DB 60 LNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPNCKH 119
QY 120 QCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQDGLL 179
DB 120 QCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQDGLL 179
QY 180 GKELGPDASEVELTRNNELIANGKRSKLEL-PVFGMEPRILYNPLOGQKCIQVTTWS 237
DB 178 SKFGLDASEGELTRNNELIANGKRSKLEL-PVFGMEPRILYNPLOGQKCIQVTTWS 232
QY 238 QCSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 297
DB 233 QCSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 292
QY 298 FTYAGCLSVKKYRKYKCGSCVDGRCCCTPQTRTVKRFCEDETFSKNVMIIQSCCKNY 357
DB 293 FTYAGCLSVKKYRKYKCGSCVDGRCCCTPQTRTVKRFCEDETFSKNVMIIQSCCKNY 352
QY 358 NCPHANEAAFPFYFLFNDIHKFRD 381
DB 353 NCPHANE-AAFPFYFLVNDIHKFRD 375

RESULT 8
US-09-582-337-2
; Sequence 2, Application US/09582337
; Patent No. 6562618
; GENERAL INFORMATION:
; APPLICANT: Japan Tobacco, Inc.
; TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor
; FILE REFERENCE: JP-009PCT
; CURRENT APPLICATION NUMBER: US/09/582,337
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: JP P1997-367699
; PRIOR FILING DATE: 1997-12-25
; PRIOR APPLICATION NUMBER: JP P1998-356183
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Rat
US-09-582-337-2

Query Match      45.0%; Score 953; DB 4; Length 347;
Best Local Similarity 46.2%; Pred. No. 2.5e-75;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSSRIARALALVLTLLHL-TRIAL-STCPAAACHCPLE-APKCAPGVGLVRDGCCKVCA 57
DB 1 MLASVAGPISLALVLLALCTRPATQDCSAQCQCAEAAPHPAGVSLVLDGCGCCRVCA 60
QY 58 KOLNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPN 117
DB 61 KOLGELCTERDPCDPHKGFLCDFGSPANRKGIVGTAK-DGAPCVFGSGVYSGSFSOSSC 119
QY 118 QHQCCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQD 177
DB 120 KYQCTCLDGAAGCVPLCSMDVRLSPDCPFPRRVKLPKGCCEEWVCDSDSIKDPMEDQD 168
```

```
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSSRIARALALVLTLLHLTRAL-STCPAAACHCPLE-APKCAPGVGLVRDGCCKVCAK 58
DB 1 MLASVAGPISLALVLLALCTRPATQDCSAQCQCAEAAPHPAGVSLVLDGCGCCRVCA 60
QY 59 OLNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPNCO 118
DB 61 QLGLCTERDPCDPHKGFLCDFGSPANRKGIVGTAK-DGAPCVFGSGVYSGSFSQSCCK 119
QY 119 HQCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQDGL 178
DB 120 YQCTCLDGAAGCVPLCSMDVRLSPDCPFPRRVKLPKGCCEEWVCDSDSIKDPMEDQD 167
QY 179 LGKELGPDASEVELTRNNELIANGKRSKLEL-PVFGMEPRILYNPLOGQKCIQVTTWS 237
DB 168 -----XDRTVVGFPAALAAAYLEDTFGDPTMM-----RANCLVQTTWS 205
QY 238 QCSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 297
DB 206 ACSKTCTGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKSKTKKSPSPVR 265
QY 298 FTYAGCLSVKKYRKYKCGSCVDGRCCCTPQTRTVKRFCEDETFSKNVMIIQSCCKNY 357
DB 266 FELSGCTSVKTYRAKFCGVCCTDGRCCCTPRTTTLFVBFKCPDGBIMKKNMFIKTCACHY 325
QY 358 NCPHANEAAFPFY--RLFNDI 376
DB 326 NCPGNDIFESLYRYKMYGDM 346

RESULT 9
US-09-292-036-3
; Sequence 3, Application US/09292036
; Patent No. 6358741
; GENERAL INFORMATION:
; APPLICANT: FIBROGEN, INC
; APPLICANT: SCHMIDT, Brian
; APPLICANT: ALLEN, Margaret
; APPLICANT: SVERDRUP, Fran
; APPLICANT: CARMICHAEL, David
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE
; FILE REFERENCE: FIBRO100-1
; CURRENT APPLICATION NUMBER: US/09/292,036
; CURRENT FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/292,036
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/187,478
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Mouse
US-09-292-036-3

Query Match      45.0%; Score 952.5; DB 4; Length 348;
Best Local Similarity 46.6%; Pred. No. 2.7e-75;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSSRIARALALVLTLLHL-TRIAL-STCPAAACHCPLE-APKCAPGVGLVRDGCCKVCA 57
DB 1 MLASVAGPISLALVLLALCTRPATQDCSAQCQCAEAAPHPAGVSLVLDGCGCCRVCA 60
QY 58 KOLNEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCYNSRIYQNGESFQPN 117
DB 61 KOLGELCTERDPCDPHKGFLCDFGSPANRKGIVGTAK-DGAPCVFGSGVYSGSFSOSSC 119
QY 118 QHQCCTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEEWVCDSDSIKDPMEDQD 177
DB 120 KYQCTCLDGAAGCVPLCSMDVRLSPDCPFPRRVKLPKGCCEEWVCDSDSIKDPMEDQD 168
```

QY 178 LLGKELGPDASEVELTRNNELIAVKGSRSLKRL-PVFGMEPRILYNPLQGGKCIQVOTTSW 236
Db 169 -----KDRTAGVPALAAVRLDTFGPDPTMM-----RANCLVQTTW 205
QY 237 SQCKTCTGISTRTVNDNPECLRVKTRICEVPRPCGQPVYSSLLKGGKCKSKTKKSPPEV 296
Db 206 SACSCTCGMISTRTVNDNTEFCRLKQSRCLQWVRPCDADLEENIKKGGKCIQVOTTSW 265
QY 297 RFTYAGCLSVKXKYPKYGSCVDGRCCTPQLTTRVGMFRCEDETFKSNVMMIQSCKN 356
Db 266 KFELSGCTSVKTYRAKFGVCTDGRCTPHTTLLPVEFKCPDGEIMKKNMFIKTCACH 325
QY 357 YNCPHANEAAFPFY--RLFNDI 376
Db 326 YNCPGNDNIFESLYRKNYGM 347
RESULT 10
US-08-468-847B-15
; Sequence 15, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-15
Query Match 44.8%; Score 948.5; DB 1; Length 348;
Best Local Similarity 46.3%; Pred. No. 6.1e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSRIARALALVVTLLHL-TRIAL-STCPAACHCPLE-APKCAPGVGLVRDGGCGCKVCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGQCSAQCCAAEAAPHCPAGVSLVLDGCGCRVCA 60
QY 58 KQINEDCSQTCDDHTKLECNFGASSTALKGICRAQSGRCEYNSRIYONGSFQNC 117
Db 61 KQIGELCTERDPCDPKHLGFLCDFGSPANKIGVCTAK-DGAPCVFGGSVYRSGSFQSC 119

QY 118 QHQTCTIDGAVGCIPLCPQELSLPNLGNPNRPLVKVTCQCEEWVCDDESINKDMEDQDQ 177
Db 120 KYQCTCLDGAAGCVPLCQMDVRLPSPDCPFPRRVKLPKCKCKEWVCDPEP----- 168
QY 178 LLGKELGPDASEVELTRNNELIAVKGSRSLKRL-PVFGMEPRILYNPLQGGKCIQVOTTSW 236
Db 169 -----KDRTAGVPALAAVRLDTFGPDPTMM-----RANCLVQTTW 205
QY 237 SQCKTCTGISTRTVNDNPECLRVKTRICEVPRPCGQPVYSSLLKGGKCKSKTKKSPPEV 296
Db 206 SACSCTCGMISTRTVNDNTEFCRLKQSRCLQWVRPCDADLEENIKKGGKCIQVOTTSW 265
QY 297 RFTYAGCLSVKXKYPKYGSCVDGRCCTPQLTTRVGMFRCEDETFKSNVMMIQSCKN 356
Db 266 KFELSGCTSVKTYRAKFGVCTDGRCTPHTTLLPVEFKCPDGEIMKKNMFIKTCACH 325
QY 357 YNCPHANEAAFPFY--RLFNDI 376
Db 326 YNCPGNDNIFESLYRKNYGM 347
RESULT 11
US-09-142-569-6
; Sequence 6, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: "Fiepl2 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-142-569-6
Query Match 44.8%; Score 948.5; DB 4; Length 348;
Best Local Similarity 46.3%; Pred. No. 6.1e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSRIARALALVVTLLHL-TRIAL-STCPAACHCPLE-APKCAPGVGLVRDGGCGCKVCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGQCSAQCCAAEAAPHCPAGVSLVLDGCGCRVCA 60

QY 58 KQNECSKTOPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCYNSRIYQNGESFPQNC 117
 DB 61 KQGLCTERDPCDPHKGFLFCDFGSPANKIGVCTAK-DGAPCVFGSVYRSGESFQSSC 119
 QY 118 HQCTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVVGQCCSEWVCDSDSIKDPMDQDG 177
 DB 120 KYQCTCLDGAAGVCLSCMDVRLSPDCPPRRVKLPGKCKEWCDEP----- 168
 QY 178 LLGKELGFDASEVELTRNNELIAGKRSKLRL-PVFGMEPRILYNPLQKQKIVQTTWS 236
 DB 169 -----KQRTAVGPALAAAYRLBDETFGPDPTM-----RANCLVQTTWS 205
 QY 237 SOCSKTGCTGISTRTVNDNPECLVKETRICVEPRGQPVVSSLKGGKCKSKTKSPVP 296
 DB 206 SACSCTCGMISTRTVNDNFTCLRKQSLCWRPCDADLEENIKGKKCIRTPKIAKPV 265
 QY 297 RFTVAGLSVKYKPYKCGSDGRCCTPOLTRTVKMRFCBDEGETFSKNVMMIQSKCN 356
 DB 266 KFELSGCTSVKTYRAKFCGVCCTGRCCTPHRTTLPVEFKCPDGEIMKKNMFIKTACH 325
 QY 357 YNCPHANEAAFPF--RLENDI 376
 DB 326 YNCPGDNDIFESLYRKMIGDM 347

RESULT 12

US-09-187-478-2

; Sequence 2, Application US/09187478

; Patent No. 6348329

; GENERAL INFORMATION:

; APPLICANT: Schmidt, Brian F.

; APPLICANT: Allen, Margaret L.

; TITLE OF INVENTION: Connective Tissue Growth (CTGF) And Methods Of Use

; FILE REFERENCE: 08766/004001

; CURRENT APPLICATION NUMBER: US/09/187,478

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 2

; TYPE: PRT

; ORGANISM: Rat

; ORGANISM: No. 6348329mal Rate Kidney Fibroblast

US-09-187-478-2

Query Match 44.8%; Score 948; DB 4; Length 347;
 Best Local Similarity 46.3%; Pred. No. 6.8e-75;
 Matches 177; Conservative 58; Mismatches 105; Indels 42; Gaps 9;

QY 1 MSSRIARALAVVTLHLTRLAL-STCPAAACHCPLA-APKCAPGVGLVRDGGCCCKYCAK 58
 DB 1 MLASVAGPVSLVALVLLCTRTATGQDCSAQCCAAEAPRCAPAGVSLVLDGCGCCRVCAK 60
 QY 59 QLNEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCYNSRIYQNGESFPQNC 118
 DB 61 QLGELCTERDPCDPHKGFLFCDFGSPANKIGVCTAK-DGAPCVFGSVYRSGESFQSSCK 119
 QY 119 HQCTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVVGQCCSEWVCDSDSIKDPMDQDG 178
 DB 120 KYQCTCLDGAAGVCLSCMDVRLSPDCPPRRVKLPGKCKEWCDEP----- 167
 QY 179 LGKELGFDASEVELTRNNELIAGKRSKLRL-PVFGMEPRILYNPLQKQKIVQTTWS 237
 DB 168 -----KQRTAVGPALAAAYRLBDETFGPDPTM-----RANCLVQTTWS 205
 QY 238 QCSKTGCTGISTRTVNDNPECLVKETRICVEPRGQPVVSSLKGGKCKSKTKSPVP 297
 DB 206 ACSCTCGMISTRTVNDNFTCLRKQSLCWRPCDADLEENIKGKKCIRTPKIAKPV 265
 QY 298 FTYAGLSVKYKPYKCGSDGRCCTPOLTRTVKMRFCBDEGETFSKNVMMIQSKCN 357
 DB 266 FELSGCTSVKTYRAKFCGVCCTGRCCTPHRTTLPVEFKCPDGEIMKKNMFIKTACH 325

RESULT 14

US-08-167-628-2

; Sequence 2, Application US/08167628

; Patent No. 5408040

; GENERAL INFORMATION:

; APPLICANT: Grotendorst, Gary R.

; APPLICANT: Bradham Jr., Douglas M.,

; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR

; NUMBER OF SEQUENCES: 2

QY 358 NCPHANEAAFP--FYR-LFNDI 376
 DB 326 NCPGDND-IFPCMYRKMIGDM 346

RESULT 13

US-09-292-036-2

; Sequence 2, Application US/09292036

; Patent No. 6358741

; GENERAL INFORMATION:

; APPLICANT: FIBROGEN, INC

; APPLICANT: SCHMIDT, Brian

; APPLICANT: ALLEN, Margaret

; APPLICANT: SVERDRUP, Fran

; APPLICANT: CARMICHAEL, David

; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE

; FILE REFERENCE: FIB01100-1

; CURRENT APPLICATION NUMBER: US/09/292,036

; CURRENT FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/292,036

; PRIOR FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/187,478

; PRIOR FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2

; LENGTH: 347

; TYPE: PRT

; ORGANISM: Rat

US-09-292-036-2

Query Match 44.8%; Score 948; DB 4; Length 347;

Best Local Similarity 46.3%; Pred. No. 6.8e-75;

Matches 177; Conservative 58; Mismatches 105; Indels 42; Gaps 9;

QY 1 MSSRIARALAVVTLHLTRLAL-STCPAAACHCPLA-APKCAPGVGLVRDGGCCCKYCAK 58
 DB 1 MLASVAGPVSLVALVLLCTRTATGQDCSAQCCAAEAPRCAPAGVSLVLDGCGCCRVCAK 60
 QY 59 QLNEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCYNSRIYQNGESFPQNC 118
 DB 61 QLGELCTERDPCDPHKGFLFCDFGSPANKIGVCTAK-DGAPCVFGSVYRSGESFQSSCK 119
 QY 119 HQCTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVVGQCCSEWVCDSDSIKDPMDQDG 178
 DB 120 KYQCTCLDGAAGVCLSCMDVRLSPDCPPRRVKLPGKCKEWCDEP----- 167
 QY 179 LGKELGFDASEVELTRNNELIAGKRSKLRL-PVFGMEPRILYNPLQKQKIVQTTWS 237
 DB 168 -----KQRTAVGPALAAAYRLBDETFGPDPTM-----RANCLVQTTWS 205
 QY 238 QCSKTGCTGISTRTVNDNPECLVKETRICVEPRGQPVVSSLKGGKCKSKTKSPVP 297
 DB 206 ACSCTCGMISTRTVNDNFTCLRKQSLCWRPCDADLEENIKGKKCIRTPKIAKPV 265
 QY 298 FTYAGLSVKYKPYKCGSDGRCCTPOLTRTVKMRFCBDEGETFSKNVMMIQSKCN 357
 DB 266 FELSGCTSVKTYRAKFCGVCCTGRCCTPHRTTLPVEFKCPDGEIMKKNMFIKTACH 325

QY 358 NCPHANEAAFP--FYR-LFNDI 376

DB 326 NCPGDND-IFPCMYRKMIGDM 346

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,628
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wetherell, Jr. Ph.D., John W.
 ; REGISTRATION NUMBER: 31,678
 ; REFERENCE/DOCKET NUMBER: PD-1294
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-455-5100
 ; TELEFAX: 619-455-5110
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 349 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-167-628-2

Query Match 44.8%; Score 947; DB 1; Length 349;
 Best Local Similarity 46.6%; Pred. No. 8.3e-75;
 Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;
 QY 9 LALVTLHLTLAL-STCPAACHCPLE-APKAPGVGLVRDGGCCCKVCAKQLEDGSK 66
 Db 11 VAFVLLALCSPPAVGQNGSGPCFDEPAPCPAGVSLVLDGGCCCRVCAKQLELCTE 70
 QY 67 TOPCDHTKLENGFASSTALXGICRAOSEGRPCYSNRIYONGESFQPNQHOCTCIDG 126
 Db 71 RDCPDHKLGLFCDFGSPANKIGVCTAK-DGAPCIFGGTVYRSGESFQSSCKYQCTCLDG 129
 QY 127 AVGCIPPLCPQELSLPNLGNPRLVKVTGQCEEWVDEDSIKOPMEDDGLLKGELGFD 186
 Db 130 AVGCMLCSMDVRLSPDCPFPRRVKLPKCCCEWVCDPE-----KDQ----- 172
 QY 187 ASEVELTRNNELIAYGKGRSLKRL-PVFGMEPRILYNPLQGGKCIIVQTTWSQCKTCGT 245
 Db 173 -----TVVGPALAAVRLDTFGDPDTMI-----RANCLVQTTENSACSKTCGM 215
 QY 246 GISTRVTNDNPECLRVKTRICEVPCQGVYSSLLKGGKCKTKSPVPVFTYAGCLS 305
 Db 216 GISTRVTNDNASCRLEKQSRLCWVRPCEADLEENIKGKKCIRTPTKISPKIFELSGCTS 275
 QY 306 VKYRPKYCGSDGRCCTPQLTRVTKMFRCEDETFESKNVMMIQSKCNKYNCPHANE 365
 Db 276 MKTYRAKFCGVCTDGRCCCTPHRTTLTPVEFKPCDGEVKKMMFMFKTCAHCYNCPGDNDI 335
 QY 366 APPFY--RLFNDI 376
 Db 336 FESLYRKMYGDM 348

RESULT 15
 US-08-386-680-2
 ; Sequence 2, Application US/08386680
 ; Patent No. 5585270
 ; GENERAL INFORMATION:
 ; APPLICANT: Grotendorst, Gary R.

; APPLICANT: Bradham Jr., Douglas M.,
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/386,680
 ; FILING DATE: 10-FEB-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wetherell, Jr. Ph.D., John W.
 ; REGISTRATION NUMBER: 31,678
 ; REFERENCE/DOCKET NUMBER: PD-1294
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-455-5100
 ; TELEFAX: 619-455-5110
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 349 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-386-680-2

Query Match 44.8%; Score 947; DB 1; Length 349;
 Best Local Similarity 46.6%; Pred. No. 8.3e-75;
 Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;
 QY 9 LALVTLHLTLAL-STCPAACHCPLE-APKAPGVGLVRDGGCCCKVCAKQLEDGSK 66
 Db 11 VAFVLLALCSPPAVGQNGSGPCFDEPAPCPAGVSLVLDGGCCCRVCAKQLELCTE 70
 QY 67 TOPCDHTKLENGFASSTALXGICRAOSEGRPCYSNRIYONGESFQPNQHOCTCIDG 126
 Db 71 RDCPDHKLGLFCDFGSPANKIGVCTAK-DGAPCIFGGTVYRSGESFQSSCKYQCTCLDG 129
 QY 127 AVGCIPPLCPQELSLPNLGNPRLVKVTGQCEEWVDEDSIKOPMEDDGLLKGELGFD 186
 Db 130 AVGCMLCSMDVRLSPDCPFPRRVKLPKCCCEWVCDPE-----KDQ----- 172
 QY 187 ASEVELTRNNELIAYGKGRSLKRL-PVFGMEPRILYNPLQGGKCIIVQTTWSQCKTCGT 245
 Db 173 -----TVVGPALAAVRLDTFGDPDTMI-----RANCLVQTTENSACSKTCGM 215
 QY 246 GISTRVTNDNPECLRVKTRICEVPCQGVYSSLLKGGKCKTKSPVPVFTYAGCLS 305
 Db 216 GISTRVTNDNASCRLEKQSRLCWVRPCEADLEENIKGKKCIRTPTKISPKIFELSGCTS 275
 QY 306 VKYRPKYCGSDGRCCTPQLTRVTKMFRCEDETFESKNVMMIQSKCNKYNCPHANE 365
 Db 276 MKTYRAKFCGVCTDGRCCCTPHRTTLTPVEFKPCDGEVKKMMFMFKTCAHCYNCPGDNDI 335
 QY 366 APPFY--RLFNDI 376
 Db 336 FESLYRKMYGDM 348

Search completed: April 22, 2004, 18:00:06

Job time : 17.5434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:58:40 ; Search time 38.6013 Seconds
(without alignments)
2728.846 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALAVTLLHLTR.....ANEAPPPYELFNDHKFRD 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1133595 seqs, 276475211 residues

Total number of hits satisfying chosen parameters: 1133595

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	381	13	US-10-053-753-4
2	2116	100.0	381	15	US-10-182-432-4
3	2106	99.5	381	10	US-09-901-910-2
4	2106	99.5	381	14	US-10-294-796-2
5	2106	99.5	381	14	US-10-394-015-5
6	2106	99.5	381	15	US-10-099-322-42
7	2106	99.5	381	15	US-10-044-564-42
8	2106	99.5	381	16	US-10-381-644-2
9	2106	99.5	381	16	US-10-464-368-61
10	2106	99.5	455	9	US-09-925-301-1432
11	2103	99.4	381	15	US-10-099-322-43
12	2103	99.4	381	15	US-10-044-564-43
13	2100	99.2	381	15	US-10-099-322-44
14	2100	99.2	381	15	US-10-044-564-44
15	2098	99.1	381	14	US-10-205-823-84

Sequence 2, Appli
Sequence 41, Appli
Sequence 2, Appli
Sequence 41, Appli
Sequence 12, Appli
Sequence 7, Appli
Sequence 11, Appli
Sequence 2, Appli
Sequence 45, Appli
Sequence 2, Appli
Sequence 45, Appli
Sequence 60, Appli
Sequence 62, Appli
Sequence 13, Appli
Sequence 63, Appli
Sequence 7, Appli
Sequence 56, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 15, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 6, Appli
Sequence 54, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 26, Appli
Sequence 8, Appli
Sequence 173, App

16 2098 99.1 381 15 US-10-039-322-2
17 2098 99.1 381 15 US-10-099-322-41
18 2098 99.1 381 15 US-10-044-564-2
19 2098 99.1 381 15 US-10-044-564-41
20 1971.5 93.2 374 9 US-09-853-625B-12
21 1971.5 93.2 375 10 US-09-901-910-7
22 1929 91.2 379 9 US-09-853-625B-11
23 1929 91.2 379 13 US-10-053-753-2
24 1929 91.2 379 15 US-10-099-322-45
25 1929 91.2 379 15 US-10-182-432-2
26 1929 91.2 379 15 US-10-044-564-45
27 1929 91.2 379 16 US-10-464-368-60
28 1928 91.1 379 16 US-10-464-368-62
29 1695 80.1 375 9 US-09-853-625B-13
30 1663.5 78.6 375 16 US-10-464-368-63
31 955 45.1 347 14 US-10-245-977-7
32 955 45.1 347 16 US-10-464-368-56
33 953 45.0 347 14 US-10-390-986-2
34 952.5 45.0 348 13 US-10-101-040-3
35 948.5 44.8 348 9 US-09-853-625B-15
36 948.5 44.8 348 13 US-10-053-753-6
37 948.5 44.8 348 14 US-10-245-977-8
38 948.5 44.8 348 15 US-10-182-432-6
39 948.5 44.8 348 16 US-10-464-368-54
40 948 44.8 347 13 US-10-080-173-2
41 948 44.8 347 13 US-10-101-040-2
42 947 44.8 349 13 US-10-101-040-4
43 947 44.8 349 13 US-10-011-859-26
44 947 44.8 349 13 US-10-053-753-8
45 947 44.8 349 14 US-10-060-036-173

ALIGNMENTS

RESULT 1
US-10-053-753-4
; Sequence 4, Application US/10053753
; Publication NO. US20020150986A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/053,753
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein

```
;
; NAME/KEY: misc feature
; OTHER INFORMATION: "Human Cyt61 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-053-753-4

Query Match      100.0%; Score 2116; DB 13; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.4e-172;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180
QY 181 KELGFDASEVELTRNNELIAGKSLKRLPVFGMEPRILYNPLOGQKCIQVOTTSWSQCS 240
DB 181 KELGFDASEVELTRNNELIAGKSLKRLPVFGMEPRILYNPLOGQKCIQVOTTSWSQCS 240
QY 241 KTCGTGISTRVTNDNPECLVKETRICVVRPGQGPVYSSLLKGGKCKSKTKKSPVPVFTY 300
DB 241 KTCGTGISTRVTNDNPECLVKETRICVVRPGQGPVYSSLLKGGKCKSKTKKSPVPVFTY 300
QY 301 AGCLSVKKYRKYCGSCVDGRCTPQLTTRTVKMRFRCEBGETTFSKNVMMIQSKCNYNCP 360
DB 301 AGCLSVKKYRKYCGSCVDGRCTPQLTTRTVKMRFRCEBGETTFSKNVMMIQSKCNYNCP 360
QY 361 HANEAAFPVRLFNDIHKFRD 381
DB 361 HANEAAFPVRLFNDIHKFRD 381

RESULT 2
US-10-182-432-4
; Sequence 4, Application US/10182432
; Publication No. US20040002124A1
; GENERAL INFORMATION:
; APPLICANT: LAU, Lester F., YEUNG, Cho-Yau, and GREENSPAN, Jeffrey A.
; TITLE OF INVENTION: CYR61 COMPOSITIONS AND METHODS
; FILE REFERENCE: 214448/00029
; CURRENT APPLICATION NUMBER: US/10/182,432
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-432-4

Query Match      100.0%; Score 2116; DB 15; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.4e-172;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180

RESULT 3
US-09-901-910-2
; Sequence 2, Application US/09901910
; Publication No. US20030012768A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark
; APPLICANT: Calenda Valerie
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PF126P2
; CURRENT APPLICATION NUMBER: US/09/901,910
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; PRIOR APPLICATION NUMBER: 60/217,402
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/291,642
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-901-910-2

Query Match      99.5%; Score 2106; DB 10; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.7e-171;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRPCENSRINYONGESFPNCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPGLGCPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG 180
QY 181 KELGFDASEVELTRNNELIAGKSLKRLPVFGMEPRILYNPLOGQKCIQVOTTSWSQCS 240
DB 181 KELGFDASEVELTRNNELIAGKSLKRLPVFGMEPRILYNPLOGQKCIQVOTTSWSQCS 240
QY 241 KTCGTGISTRVTNDNPECLVKETRICVVRPGQGPVYSSLLKGGKCKSKTKKSPVPVFTY 300
DB 241 KTCGTGISTRVTNDNPECLVKETRICVVRPGQGPVYSSLLKGGKCKSKTKKSPVPVFTY 300
QY 301 AGCLSVKKYRKYCGSCVDGRCTPQLTTRTVKMRFRCEBGETTFSKNVMMIQSKCNYNCP 360
```

```

Db 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFRCEDEGTFTSKNVMIIQSKCKNYP 360
Qy 361 HANEAAFPYRLFNDFHFRD 381
Db 361 HANEAAFPYRLFNDFHFRD 381

RESULT 4
US-10-294-796-2
; Sequence 2, Application US/10294796
; Publication No. US20030078391A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong et al.
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PF126P1D2
; CURRENT APPLICATION NUMBER: US/10/294,796
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-796-2

Query Match 99.5%; Score 2106; DB 14; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.7e-171;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSSRIARALALVVTLLHLTRIALSTCPAAACHPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
Db 1 MSSRIARALALVVTLLHLTRIALSTCPAAACHPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRYYQNGESFQPCNCHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRYYQNGESFQPCNCHQ 120
Qy 121 CTCIDGAVGCIPLCPQELSPLNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDDGLLG 180
Db 121 CTCIDGAVGCIPLCPQELSPLNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDDGLLG 180
Qy 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLQGGKCIQVTTTWSQCS 240
Db 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLQGGKCIQVTTTWSQCS 240
Qy 241 KTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
Db 241 KTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
Qy 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFRCEDEGTFTSKNVMIIQSKCKNYP 360
Db 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFRCEDEGTFTSKNVMIIQSKCKNYP 360
Qy 361 HANEAAFPYRLFNDFHFRD 381
Db 361 HANEAAFPYRLFNDFHFRD 381

RESULT 5
US-10-394-015-5
; Sequence 5, Application US/10394015
; Publication No. US20030180891A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Connective Tissue Growth Factor-4
; FILE REFERENCE: PF467

```

```

; CURRENT APPLICATION NUMBER: US/10/394,015
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US/09/325,019
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/088,320
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-015-5

Query Match 99.5%; Score 2106; DB 14; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.7e-171;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSSRIARALALVVTLLHLTRIALSTCPAAACHPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
Db 1 MSSRIARALALVVTLLHLTRIALSTCPAAACHPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRYYQNGESFQPCNCHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRYYQNGESFQPCNCHQ 120
Qy 121 CTCIDGAVGCIPLCPQELSPLNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDDGLLG 180
Db 121 CTCIDGAVGCIPLCPQELSPLNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDDGLLG 180
Qy 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLQGGKCIQVTTTWSQCS 240
Db 181 KELGPDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLQGGKCIQVTTTWSQCS 240
Qy 241 KTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
Db 241 KTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
Qy 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFRCEDEGTFTSKNVMIIQSKCKNYP 360
Db 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFRCEDEGTFTSKNVMIIQSKCKNYP 360
Qy 361 HANEAAFPYRLFNDFHFRD 381
Db 361 HANEAAFPYRLFNDFHFRD 381

RESULT 6
US-10-099-322-42
; Sequence 42, Application US/10099322
; Publication No. US20030215449A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/10/099,322
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11

```

NUMBER OF SEQ ID NOS: 324
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 42
 LENGTH: 381
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-099-322-42

Query Match 99.5%; Score 2106; DB 15; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.7e-171;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIIVTTTWSQCS 240
 DB 181 KELGFDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIIVTTTWSQCS 240
 QY 241 KTCGTGISTRVTNDNPECLVKETRICVRCQGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 DB 241 KTCGTGISTRVTNDNPECLVKETRICVRCQGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 QY 301 AGCLSVKRYPKYCGSCVDGRCCTPOLTRTVKMRFRCEDETFKSNVMMIOCKKNYNC 360
 DB 301 AGCLSVKRYPKYCGSCVDGRCCTPOLTRTVKMRFRCEDETFKSNVMMIOCKKNYNC 360
 QY 361 HANEAAFPYRLFNDFHFRD 381
 DB 361 HANEAAFPYRLFNDFHFRD 381

RESULT 7

US-10-044-564-42
 ; Sequence 42, Application US/10044564
 ; Publication NO. US20040018196A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes et al.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-240
 ; CURRENT APPLICATION NUMBER: US/10/044,564
 ; PRIOR FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 60/261,014
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,018
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/318,410
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/261,013
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,026
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,029
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/313,170
 ; PRIOR FILING DATE: 2001-08-17
 ; NUMBER OF SEQ ID NOS: 306
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 42
 ; LENGTH: 381
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-044-564-42

Query Match 99.5%; Score 2106; DB 15; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.7e-171;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180
 QY 181 KELGFDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIIVTTTWSQCS 240
 DB 181 KELGFDASEVELTRNNELIANGKSLKRLPVFGMEPRILYNPLOGQKCIIVTTTWSQCS 240
 QY 241 KTCGTGISTRVTNDNPECLVKETRICVRCQGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 DB 241 KTCGTGISTRVTNDNPECLVKETRICVRCQGPVYSSLLKGGKCKSKTKSPVPVFTY 300
 QY 301 AGCLSVKRYPKYCGSCVDGRCCTPOLTRTVKMRFRCEDETFKSNVMMIOCKKNYNC 360
 DB 301 AGCLSVKRYPKYCGSCVDGRCCTPOLTRTVKMRFRCEDETFKSNVMMIOCKKNYNC 360
 QY 361 HANEAAFPYRLFNDFHFRD 381
 DB 361 HANEAAFPYRLFNDFHFRD 381

RESULT 8

US-10-381-644-2
 ; Sequence 2, Application US/10381644
 ; Publication NO. US20040023910A1
 ; GENERAL INFORMATION:
 ; APPLICANT: American Home Products Corporation
 ; APPLICANT: Zhang, Zhiming
 ; APPLICANT: Sampath, Deepak
 ; APPLICANT: Zhu, Yuan
 ; APPLICANT: Winnekar, Richard
 ; TITLE OF INVENTION: Use of Cyr61 in the treatment and
 ; FILE REFERENCE: AM100352
 ; CURRENT APPLICATION NUMBER: US/10/381,644
 ; PRIOR FILING DATE: 2003-03-26
 ; PRIOR APPLICATION NUMBER: 60/236,887
 ; PRIOR FILING DATE: 2000-09-29
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 381
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-381-644-2

Query Match 99.5%; Score 2106; DB 16; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.7e-171;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLNCFNGFASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGCCCEEWVCDSDSIKDPMEDQDGLLG 180

181 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
181 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
241 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 300
241 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 300
301 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 360
301 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 360
361 HANEAAFPFYRLFNDIHKFRD 381
361 HANEAAFPFYRLFNDIHKFRD 381

RESULT 9
US-10-464-368-61
; Sequence 61, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 61
; LENGTH: 381
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-61

Query Match 99.5%; Score 2106; DB 16; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.7e-171;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLTRALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVYNSRIYQNGESFPNCQHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVYNSRIYQNGESFPNCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDQDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDQDGLLG 180
QY 181 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
DB 181 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
QY 241 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 300
DB 241 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 300
QY 301 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 360
DB 301 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 360
QY 361 HANEAAFPFYRLFNDIHKFRD 381
DB 361 HANEAAFPFYRLFNDIHKFRD 381

RESULT 10

US-09-925-301-1432
; Sequence 1432, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1432
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1432

Query Match 99.5%; Score 2106; DB 9; Length 455;
Best Local Similarity 99.5%; Pred. No. 2.1e-171;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLTRALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 75 MSSRIARALALVVTLLHLTRALSTCPAAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 134
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVYNSRIYQNGESFPNCQHQ 120
DB 135 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVYNSRIYQNGESFPNCQHQ 194
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDQDGLLG 180
DB 195 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVTGQCCCEEWVCDSDSIKDPMEDQDGLLG 254
QY 181 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
DB 255 KELGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 314
QY 241 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 300
DB 315 KTCGTGISTRVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKTKKSPPEVRFTY 374
QY 301 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 360
DB 375 AGCLSVKRYRKYKCGSCVDGRCCCTPQLTRTVKMRFRCEDEGETFSKNVMIQSKCKNYNCP 434
QY 361 HANEAAFPFYRLFNDIHKFRD 381
DB 435 HANEAAFPFYRLFNDIHKFRD 455

RESULT 11
US-10-099-322-43
; Sequence 43, Application US/10099322
; Publication No. US2003021549A1
; GENERAL INFORMATION:
; APPLICANT: Vezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/10/099,322
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026

```
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-43

Query Match          99.4%; Score 2103; DB 15; Length 381;
Best Local Similarity 99.2%; Pred. No. 3.1e-171;
Matches 378; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLKAPKAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLKAPKAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCYNRIYQNGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCYNRIYQNGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDIKDPMEDQDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDIKDPMEDQDGLLG 180
QY 181 KELGPDASEVELTRNNELIIVGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
DB 181 KELGPDASEVELTRNNELIIVGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
QY 241 KTCGTGISTRVNDNPECLVKETRICVRCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
DB 241 KTCGTGISTRVNDNPECLVKETRICVRCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
QY 301 AGCLSVKKYRPKYCGSVDGRCCTPQLTRTVKMRFRCEDETFSSKNVMMIQSKCKNKNCP 360
DB 301 AGCLSVKKYRPKYCGSVDGRCCTPQLTRTVKMRFRCEDETFSSKNVMMIQSKCKNKNCP 360
QY 361 HANEAAFFPYRLFNDFHFRD 381
DB 361 HANEAAFFPYRLFNDFHFRD 381

RESULT 13
US-10-099-322-44
; Sequence 44, Application US/10099322
; Publication No. US20030215449A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/10/099,322
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-44

; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-43

Query Match          99.4%; Score 2103; DB 15; Length 381;
Best Local Similarity 99.2%; Pred. No. 3.1e-171;
Matches 378; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLKAPKAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLKAPKAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCYNRIYQNGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALXGICRAQSEGRPCYNRIYQNGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDIKDPMEDQDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDIKDPMEDQDGLLG 180
QY 181 KELGPDASEVELTRNNELIIVGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
DB 181 KELGPDASEVELTRNNELIIVGKGRSLKRLPVFGMEPRILYNPLQGGKCIQVTTSSQCS 240
QY 241 KTCGTGISTRVNDNPECLVKETRICVRCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
DB 241 KTCGTGISTRVNDNPECLVKETRICVRCGQPVYSSLLKGGKCKTKKSPPEPVFTY 300
QY 301 AGCLSVKKYRPKYCGSVDGRCCTPQLTRTVKMRFRCEDETFSSKNVMMIQSKCKNKNCP 360
DB 301 AGCLSVKKYRPKYCGSVDGRCCTPQLTRTVKMRFRCEDETFSSKNVMMIQSKCKNKNCP 360
QY 361 HANEAAFFPYRLFNDFHFRD 381
DB 361 HANEAAFFPYRLFNDFHFRD 381

RESULT 12
US-10-044-564-43
; Sequence 43, Application US/10044564
; Publication No. US20040018196A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240
; CURRENT APPLICATION NUMBER: US/10/044,564
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 306
```


QY	61	NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCBYNSRIYONGESFQPNCHQ	120
Db	61	NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCBYNSRIYONGESFQPNCKHQ	120
QY	121	CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLG	180
Db	121	CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVTGCCCEWVCDSDSIKDPMEDDGLLG	180
QY	181	KELGFDASEVELTRNNELIAGVKGSLKRLPVFGMEPRILYNPLQGQKCIYQTTWSQCS	240
Db	181	KELGFDASEVELTRNNELIAGVKGSSLKRIPVFGMEPRIRYNPLQGQKCIYQTTWSQCS	240
QY	241	KTCGTGISTRVTNDNPECELAKETRICEVRPGQPVYSSLLKXKCKSKTKSPVPRFTY	300
Db	241	KTCGTGISTRVTNDNPECLRVKETRICEVRPGQPVYSSLLKXKCKSKTKSPVPRFTY	300
QY	301	AGCLSVKKYRPKYCGSCVDGRCTTQLTRTVQMRFCEDGETFSKNVMMIOSCKCNYNCP	360
Db	301	AGCLSVKKYRPKYCGSCVDGRCTTQLTRTVQMRFCEDGETFSKNVMMIOSCKCNYNCP	360
QY	361	HANEAAFPFYRLFNDIHKFRD	381
Db	361	HANEAAFPFYRLFNDIHKFRD	381

Search completed: April 22, 2004, 18:08:17
Job time : 39.6013 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:53:39 ; Search time 16.5434 Seconds
(without alignments)
2215.320 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALAVVTLHLTR.....ANEAAPFYFLFNDIHKFRD 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1929	91.2	379	2 A35669	Gene CYR61 protein
2	1695	80.1	375	2 A41428	CEP-10 protein pre
3	952.5	45.0	348	2 A40578	Beta IG-W2 protein
4	947	44.8	349	2 A40551	connective tissue
5	851.5	40.2	351	2 S20078	NOV protein - chic
6	827.5	39.1	357	2 I38069	gene novH protein
7	174	8.2	1111	2 T26972	hypothetical prote
8	173.5	8.2	1620	2 T27283	hypothetical prote
9	157	7.4	1700	2 S08167	Balbani ring 3 pr
10	153	7.2	1574	2 T13954	MSGF6 protein - ra
11	152	7.2	5376	2 T42215	zonadhesin mouse
12	150.5	7.1	1178	1 A39804	thrombospondin pre
13	150	7.1	837	2 A42112	mucin-like peptide
14	147.5	7.0	1170	2 A53612	laminin B1k chain
15	146	6.9	1964	2 T93059	notch4 - mouse
16	145	6.9	601	2 T22025	hypothetical prote
17	145	6.9	601	2 D89711	protein F40E10.4 l
18	141.5	6.7	13288	2 T03099	mucin, submaxillar
19	141	6.7	1042	2 A57534	mucin 5AC (clone L
20	141	6.7	3020	2 A43932	mucin 2 precursor,
21	140.5	6.6	1025	2 T42626	secreted leucine-r
22	140.5	6.6	1034	2 J25598	mucin - rat
23	140	6.6	4135	2 T42629	tenascin-X - bovin
24	139.5	6.6	1101	2 T16840	hypothetical prote
25	139	6.6	2437	2 S42612	transmembrane prot
26	138.5	6.5	2703	1 A24420	notch protein - fr
27	138	6.5	1056	2 A53767	mucin MUC5B, trach
28	137.5	6.5	1480	2 A35665	slit protein 1 pre
29	137	6.5	2531	2 S18188	notch protein homo

30 135 6.4 1168 2 I56985 kalinin B1 - mouse
31 135 6.4 1170 2 A40558 thrombospondin 1 p
32 134.5 6.4 330 2 T25169 hypothetical prote
33 134.5 6.4 406 2 T03070 probable tenascin
34 134 6.3 1959 1 AGPT agrin - rat
35 134 6.3 3106 1 S53868 laminin alpha-2 ch
36 134 6.3 3672 2 T23433 hypothetical prote
37 134 6.3 3704 2 T37316 probable laminin a
38 133.5 6.3 1955 1 AGCH agrin precursor -
39 133 6.3 2813 1 VMHU von Willebrand fac
40 132.5 6.3 2555 2 A40043 notch protein homo
41 132 6.2 251 2 A55035 cysteine-rich prot
42 132 6.2 1469 2 B36665 slit protein 2 pre
43 132 6.2 1531 2 T42218 slit-1 protein hom
44 132 6.2 2823 2 T23064 hypothetical prote
45 132 6.2 2823 2 F87908 protein T22A3.8 l

ALIGNMENTS

RESULT 1

A35669
Gene CYR61 protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Sep-1990 #sequence revision 18-Nov-1992 #text_change 05-Nov-1999
C;Accession: A35669, I48319, S16446
R;O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
Mol. Cell. Biol. 10, 3569-3577, 1990
A;Title: Expression of cyr61, a growth factor-inducible immediate-early gene.
A;Reference number: A35669; MUID:90287146; PMID:2355916
A;Accession: A35669
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-379 <OAB>
A;Cross-references: GB:M32490; NID:9192909; PIDN:AAA37512.1; PID:9309206
A;Note: the authors translated the codon GAT for residue 337 as Gln
R;Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
Nucleic Acids Res. 19, 3261-3267, 1991
A;Title: Promoter function and structure of the growth factor-inducible immediate early
A;Reference number: I48319, MUID:91288203; PMID:2062642
A;Accession: I48319
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <RES>
A;Cross-references: EMBL:X56790; NID:950632; PIDN:CAA40109.1; PID:950633
A;Note: the authors did not translate the codon for residue 108
A;Note: the authors translated the codon GAT for residue 337 as Gln
C;Genetics:
A;Gene: CYR61
A;Introns: 21/3; 93/1; 208/1; 279/3
C;Superfamily: von Willebrand factor type C repeat homology
P;99-166/Domain: von Willebrand factor type C repeat homology <VMC>

Query Match 91.2%; Score 1929; DB 2; Length 379;
Best Local Similarity 90.9%; Pred. No. 1.3e-134; Indels 6; Gaps 2;
Matches 348; Conservative 10; Mismatches 19;

Qy 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLCAPKCAPGVGLVRDGCCKVC AKOL 60
Db 1 MSSSTFRTTLLAVVTLHLTRLALSTCPAACHCPLCAPKCAPGVGLVRDGCCKVC AKOL 60
Qy 61 NECDKSTPCDHTKGLFCNFGASSTALKGICRAQSGRCEYNRSRIYQNGESFQPNCKHQ 120
Db 61 NECDKSTPCDHTKGLFCNFGASSTALKGICRAQSGRCEYNRSRIYQNGESFQPNCKHQ 120
Qy 121 CTCIDGAVGICPLCPQELSIPNLGCPNRLVKVTGCCCEEWVCDSDSIPKDMEDQDGLLG 180
Db 121 CTCIDGAVGICPLCPQELSIPNLGCPNRLVKVTGCCCEEWVCDSDSIPKDMEDQDGLLG 180
Qy 181 KELGFDPASEVELTRNNELIYAVKGRSLKELFVGEPRILYINPL--QGOKCIVQTTWSQ 238
Db 179 --LGLDASEVELTRNNELIYAVKGRSLKELFVGEPRILYINPL--QGOKCIVQTTWSQ 236

QY 239 CSKTCGTGISTRVTNDNPECLRVKTRICEVVRPCGQPVYSSLLKGGKCSKTKKSPPEVRF 298
Db 237 CSKSGTGTGISTRVTNDNPECLRVKTRICEVVRPCGQPVYSSLLKGGKCSKTKKSPPEVRF 296
QY 299 TVAGCLSVKKYRPKYKCGSCVDGRCTPOLTTRTVKMRFCEDGETFSKNVMMIOSCKNYN 358
Db 297 TVAGSSVKKYRPKYKCGSCVDGRCTPOLTTRTVKMRFCEDGETFSKNVMMIOSCKNYN 356
QY 359 CPANEAAPFPYRLFNDFHIFKFD 381
Db 357 CPHNEASFRLYSLEFNDFHIFKFD 379

RESULT 2

A41428
CSP-10 protein precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
C:Accession: A41428
R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
A:Title: Identification of a phorbol ester-repressible v-src-inducible gene.
A:Reference number: A41428; MUID:89145206; PMID:2537491
A:Accession: A41428
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-375 <SIM>
A:Cross-references: GB:J04496; NID:g211435; PIDN:AAA48661.1; PID:g211436

Query Match 80.1%; Score 1695; DB 2; Length 375;
Best Local Similarity 81.0%; Pred. No. 2.1e-117;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSRIARALALVVTLLHL-TRLAL-STCPAACHCPLLE-APKCAPGVGLVRDGGCGCKVCAQ 59
Db 1 MGSAGARP-ALAAALLCLARLALGSPCPAVCCQCPAAAPQCAPGVGLVRDGGCGCKVCAQ 59
QY 60 LNEDESKTOPCDHTKGLCNFGASSTALKGICRAOSEGRPCENYRIYONGESFQPNCOH 119
Db 60 LNEDESKTOPCDHTKGLCNFGASSTALKGICRAOSEGRPCENYRIYONGESFQPNCKH 119
QY 120 QCTCIDGAVGCIPLCPQELSLNLCGPNRLVKVTCQCEBWCDEDSIKDPMEDQGLL 179
Db 120 QCTCIDGAVGCIPLCPQELSLNLCGPNRLVKVTCQCEBWCDEDS--KDALBELEGFF 177
QY 180 GKELGFDASEVELTRNNELIAVGKRSKRLPLVFGMEP--RILYNPLQGGKCIIVQTTWS 237
Db 178 SKFGLDASEGELTRNNELIAVGK-GLXPLVFGSEFQSRAFNP-----KCIIVQTTWS 232
QY 238 QCSKTCGTGISTRVTNDNPECLRVKTRICEVVRPCGQPVYSSLLKGGKCSKTKKSPPEVRF 297
Db 233 QCSKTCGTGISTRVTNDNPECLRVKTRICEVVRPCGQPVYSSLLKGGKCSKTKKSPPEVRF 292
QY 298 FTYAGCLSVKKYRPKYKCGSCVDGRCTPOLTTRTVKMRFCEDGETFSKNVMMIOSCKNY 357
Db 293 FTYAGSSVKKYRPKYKCGSCVDGRCTPOLTTRTVKMRFCEDGETFSKNVMMIOSCKNY 352
QY 358 NCPHANEAAFPYRLFNDFHIFKFD 381
Db 353 NCPHANE-AYPFRVLNDFHIFKFD 375

RESULT 3

A40578
beta IG-M2 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
C:Accession: A40578; A53228
R:Brunker, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
DNA Cell Biol. 10, 293-300, 1991
A:Title: Identification of a gene family regulated by transforming growth factor-beta.
A:Reference number: A40578; MUID:91229699; PMID:2029337

A:Accession: A40578
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <BRU>
A:Cross-references: GB:M80263; NID:g201945; PIDN:AAA73135.1; PID:g201946
R:Rybeck, R.P.; Macdonald-Bravo, H.; Matter, M.G.; Bravo, R.
Cell Growth Differ. 2, 225-233, 1991
A:Title: Structure, mapping, and expression of liss-12, a growth factor-inducible gene.
A:Reference number: A53228; MUID:91363290; PMID:1888698
A:Accession: A53228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160, 'K', 162-348 <RYS>
A:Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
C:Genetics:
A:Gene: fisp-12

Query Match 45.0%; Score 952.5; DB 2; Length 348;
Best Local Similarity 46.6%; Pred. No. 7.9e-63;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALALVVTLLHL-TRLAL-STCPAACHCPLLE-APKCAPGVGLVRDGGCGCKVCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCQCAEAAPHPAGVSLVLDGGCGCRVCA 60
QY 58 KOLNEDCSKTOPCDHTKGLCNFGASSTALKGICRAOSEGRPCENYRIYONGESFQPN 117
Db 61 KQGELECTERDCDHPKGLFCDFGSPANKIGVCTAK-DGAPCVFGGSYVRSGESFQSSC 119
QY 118 QHCTCIDGAVGCIPLCPQELSLNLCGPNRLVKVTCQCEBWCDEDSIKDPMEDQOG 177
Db 120 KYQCTCLDGAAGVCPVLCSDMDVRLPSPDCPPFRVRLPGKCCBEWVCDPEP----- 168
QY 178 LLGKELGFDASEVELTRNNELIAVGKRSKRLPLVFGMEPRLYNPLQGGKCIIVQTTWS 236
Db 169 -----KORTAVGDPALAAAYLEDTFGDPDTWM-----RANCLVQTTWM 205
QY 237 SQSKTCGTGISTRVTNDNPECLRVKTRICEVVRPCGQPVYSSLLKGGKCSKTKKSPPEV 296
Db 206 SACSCTGCMGISTRVTNDNTFCRLKQSLCMLVVRPCADLEENIKGKKCIRTPIAKPV 265
QY 297 RPTAGCLSVKKYRPKYKCGSCVDGRCTPOLTTRTVKMRFCEDGETFSKNVMMIOSCKN 356
Db 266 KFELSCTSVKTRAKPGVCTDGRCTPHRTITLTVFVKCPDGEIMKMMFIKTCACH 325
QY 357 YNCPHANEAAFPY--RLFNDFI 376
Db 326 YNCPGNDIFESLYRKMVYGD 347

RESULT 4

A40551
connective tissue growth factor - human
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C:Accession: A40551; S44205
R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
J. Cell Biol. 114, 1285-1294, 1991
A:Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human vas-
A:Reference number: A40551; MUID:91373462; PMID:1654338
A:Accession: A40551
A:Molecule type: mRNA
A:Residues: 1-349 <BEA>
A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AAA91279.1; PID:g180923
R:Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.; Luescher, T.F.
submitted to the EMBL Data Library, April 1994
A:Description: Differential cloning and expression of human connective tissue growth fa-
A:Reference number: S44205
A:Accession: S44205
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-349 <DEM>
A:Cross-references: EMBL:X78947; NID:g474933; PID:g474934

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20293
A;Accession: T26972
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1111 <WIL>
A;Cross-references: EMBL:AL032657; PIDN:CRAA21739.1; GSPDB:GN00019; CBSP:Y47H9C.4
A;Experimental source: clone Y47H9C
C;Genetics:
A;Gene: CBSP:Y47H9C.4
A;Map position: 1
A;Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match 8.2%; Score 174; DB 2; Length 1111;
Best Local Similarity 19.7%; Pred. No. 4.3e-05;
Matches 84; Conservative 33; Mismatches 161; Indels 148; Gaps 17;

QY 26 CPAACHCPLKLE-APKCAPGVGLYR-----DGCOCCKVCAK-QLNEDCSKTPQCDHTKGLK 78

Db 380 CSKTCCTCVRENTLMCAPNTGFCRKPFGYDNCCLACSKDSYGNCEKQAMCDMNHASEC 439

QY 79 NFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPCQHQCTC-----IDGA----- 127

Db 440 NPETGSC-----VCKPORTGKNCSEPCPL-----DFYGNCAHQCCQCNQGVGCDGADGKQ 491

QY 128 -----VGCIPL-----CPQELSLFN--LGCP-- 146

Db 492 CDGWTGHRCEHHPADTFEGANCEKCKPKGIGCDPITGECTCPAGLQGANCDIGCPGEG 551

QY 147 -----NPLVAVTQCCEWVCDSDSKDPMEDDQGLLKGELGFDASEVELT 193

Db 552 SYVPGCKLHCKVNGKCDKETGEC-----TC-----QPGFSGSDCTSTCKGKYG 596

QY 194 RNNEILAVGKRSKELPVFGMEPRILYNPLOGQKCIQVTTWSQCSKTCGTGISTRTVN 253

Db 597 ESECEL-----SCPCSASCSKQTKCLPLGTGVS 627

QY 254 DNPECLVETRICVVRPGQVPVSLKKGKCKTKSPKSPVVRTY-----AGCLS 305

Db 628 CDQCKDPNTFGLCQETVFPSCADTPKNGVCLSPGSSGSIHCEHNCPCAGSYDGGCQ 687

QY 306 VKKYRPKYCGSDVGRCTTPQLTRTV-----KVRFCEDGETFSKNVMMIQCKNY 357

Db 688 V-----C-SCADHGCDPTTGECICEPGYHGKTCSEKCPDQKYGYGCALDCPKASGS 739

QY 358 NCPHAN 363

Db 740 TCDHIN 745

RESULT 8

T27283

hypothetical protein Y64G10A.f - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T27283

R;Ainscough, R.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z20336

A;Accession: T27283

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1620 <WIL>

A;Cross-references: EMBL:AL110498; NID:61542303; PIDN:CAB54471.1; CBSP:Y64G10A.f

A;Experimental source: clone Y64G10A

C;Genetics:

A;Gene: CBSP:Y64G10A.f

A;Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7

Query Match

Best Local Similarity 24.2%; Score 173.5; DB 2; Length 1620;

Matches 92; Conservative 33; Mismatches 146; Indels 109; Gaps 21;

QY 26 CPAACHCPLKLEAP-----KCAPG-----VGLVRDGGCGCKVC AKQLNEDCSK 66

Db 1052 CKGICSCQNGATCDSVTGCECRPGRGKCKDRPCPDG--RFEGCNAICDCTTTNDTSM 1109

QY 67 TQP-----CDHTKLECNFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPCQHQCT 122

Db 1110 YNPFVARCHVTG-EGR-----CPAGWTGDCQTSCLGRHGE-----GCRHSCQ 1153

QY 123 CIDGAVGCIPLCPQELSLNLCNPNRLVAVTQCCEWVCDSDSKDPMEDDQGLLKGEL 192

Db 1154 CSNGA-----SCDRVTFG--CDQPSFGMGKNCSECEGELGWSNCKMKHCLCMKHGSGCNKE 1206

QY 183 LGFDASEVELTRNELLIAVKGSRSLKRLPVFGMEPRILYNPLOGQKCIQVOT-----TSW 236

Db 1207 NG-----DCEDIGWTGSPSLCPFGQFGNCAQRCKNKGASCDRKTGRCECLPGW 1256

QY 237 S--QCSKTCTGTGISTRTVNDNPECLVETRICVVRPGQVPVSLK-----KGKCKSKT 289

Db 1257 SGEHCKSKVSG-----HYGAKC-----EETCECENGALCDPISGHSCQCPGWRGKCKNRP 1307

QY 290 KKSPEPVRFYAGCLSVKYPKPYCG-----SC--VDGRCTTPQLTRTVMEFRCE 338

Db 1308 -----CL--KGYFGRHQSQRCSKSKSDHISGRCCQCPKGYAGHSCTELCP 1352

QY 339 DGETFSKNVMMIQCKKNYN 358

Db 1353 DG-TFGESCS--QKCDGGEN 1369

RESULT 9

S08167

Balbiani ring 3 protein - midge (Chironomus tentans)

C;Species: Chironomus tentans

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C;Accession: S08167

R;Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.

J. Mol. Biol. 211, 331-349, 1990

A;Title: the Balbiani ring 3 gene in Chironomus tentans has a diverged repetitive struc

A;Reference number: S08167; MUID:90172404; PMID:1689777

A;Accession: S08167

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-1700 <PAU>

A;Cross-references: GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058

C;Genetics:

A;Gene: BR3

A;Map position: 4

C;Superfamily: unassigned Balbiani ring proteins

Query Match

Best Local Similarity 7.4%; Score 157; DB 2; Length 1700;

Matches 88; Conservative 47; Mismatches 170; Indels 116; Gaps 21;

QY 30 CHCPLKLEAPK--CAPGVGLVRDGGC--CKVCAKQLNEDCSKTPQCDHTKGLK-----NFG 81

Db 221 CICPTAPAGGCSAPLKWDDDKSCACAPKAMEEKKEKVESGKIMNPNTCEGCAQLNCP 280

QY 82 ASSTALKGICRAQ--SEGRPCVNSRIYONGESFQPCQHQCTCI----- 124

Db 281 DNKKANKETCQCKEVKGC-----NGQVP---CKDSCSCVCPGGDKDKTCTAPQVY 330

QY 125 DGAVGCIPLCPQELSLNLCNPNRLVAVTQCCEWVCDSDSKDPMEDDQGLLKGEL 184

Db 331 DG-VACSCSCFVNMQKPADGCPRP-----QKWDKEECRCCECFVK-HDCKNGKQVMD 378

QY 185 FDASEVELTNNELLIAVKGSR-----SLKELPVFGMEPRILYNPLOGQKCIQVTTSW 236

Db 379 ETTCCCTCPDAPVCTAGKERCESCEKCKINKEPKGCAKPLVMNE-NTKCCVCVCPADK- 436

QY 237 SQCSKTCTGTGISTRTVNDNPEC-----RLVKETRICVVRP-----CQQPVYSSL 280

Db 437 QMSPGGSGSKSFNKLTCQCECDQSASKGLKKNWADTKCECQCPGMPGCGKQTWISD 496

QY 281 KKGKCKSKT-----KSGPEPVRFYAGCLSVKRYR- 311
 Db 497 KCKCSCSTITQAPQILDNTCECKCFVNLAKCKSPQWTDKCLCECSTTATC 556
 QY 312 ----KYCG---SCV-----DGRCTTQTLRTVQMRFCEDGRT-PSKNVMMIQCKKNVNC 359
 Db 557 EKGQWCEGACQICPGGDKNGKFKFDKPSCECKCKNNFTCTSPQWDDADCEC--KC 614
 QY 360 P 360
 Db 615 P 615

RESULT 10
 T13954
 MEGF6 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13954
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:98360089; PMID:9693030
 A:Accession: T13954
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1574 <NAK>
 A:Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BA032462.1; PID:G3449294
 A:Experimental source: strain Sprague-Dawley; brain
 C:Genetics:
 A:Gene: MEGF6

Query Match 7.2%; Score 153; DB 2; Length 1574;
 Best Local Similarity 21.9%; Pred. No. 0.0021;
 Matches 97; Conservative 36; Mismatches 133; Indels 176; Gaps 28;

QY 26 CPAACHCP-----LEAPKCAPGV--GLVRDGC----- 50
 Db 568 CSSPTCQNGGTCDPVLGACRCPPGVGAHCEGCPKGFYKHKRCKCHCANRGRCHRLY 627
 QY 51 GCKKCAKQL-----NEPCSKTQPCDHTKGLCNFGASSTALKGICR 92
 Db 628 GAC-LCDPGLYGRFCHLACPWAFPGCSEDLCEQ--SHTRSNPKDGSCS-----CK 678
 QY 93 AQSEGRPCVNSRIYQNGESPQNCQHOCTCIDGAVGCIPL-----CPQ 136
 Db 679 AGFOGERQAE-----ESG-FFGPGGRHRTCPQG-VACDPVSGCRCTQCPPGYQGEDCGQ 733
 QY 137 ELSLPNLG-----CNPRLVKVTGQC-CEENVCDSDSIKDPMEDQDGLGKELGFDA 187
 Db 734 ECPVTFVNGSGSCSVGAPCHRVTEGELCPGPKTGEDCGACDFGRWGLGQEI----- 789
 QY 188 SEVELTRNELLIAVGKGRSLKRLPVFGMEPRILYNPLQGGKICVQTTSSWQCSKTC--- 243
 Db 790 -----CPACEHGASC-----NPETGTCLLPFPVSGRCQDTCAGW 825
 QY 244 -GTGISTR--VTNDNPECLVKETRICVPRCGOPVYSSLKGGKCKTKKSPFVRFY 300
 Db 826 YGTGCGQIRCAANDG-HCDPTTGRCS-----APGWTGLSCQACDSDSGHWGPD----- 872
 QY 301 AGCLSVKRYKRYC-----GSC--VDGRG-CT-----PQLTRTVKMRF---RCDGETFSK 345
 Db 873 --CI-----HPCNSAGHNCDAVSLGLCLCEAGVEGPRCSQCRQGYGFSCE----- 918
 QY 346 NVMMIQCKKNYN--CPHANE 365
 Db 919 -----QKCRCHGAACDHVSGA 935

RESULT 11
 T42215
 zonadhesin - mouse

N:Alternate names: sperm-specific membrane protein
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T42215
 R:Gao, Z.; Garbers, D.L.
 J. Biol. Chem. 273, 3415-3421, 1998
 A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane protein
 A:Reference number: Z22080; MUID:98123114; PMID:9452463
 A:Accession: T42215
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-5376 <GAO>
 A:Cross-references: EMBL:U97068; NID:G3327420; PID:G3327421; PIDN:AAC26680.1
 C:Genetics:
 A:Gene: Zan
 A:Map position: 5
 C:Function:
 A:Description: functions in multiple cell adhesion processes
 A:Note: found exclusively on the apical region of the sperm head
 C:Keywords: cell adhesion

Query Match 7.2%; Score 152; DB 2; Length 5376;
 Best Local Similarity 21.0%; Pred. No. 0.0082;
 Matches 81; Conservative 39; Mismatches 138; Indels 128; Gaps 19;

QY 19 TRLALSTCPAAACHCPLEAPKCAPGVGLVRDGCCKVCAQLNEDCSKTQPCDHTKGLSC 78
 Db 4378 TKIALQ-CPAHSHTYCLPSCIPSCSNVNDRC-----ESTSLQRPSTCIEGCLC 4425
 QY 79 NFGASSTALKGICRAQ-----SEGRPCVNSRIYQNGESF-QPNQHOCTCIDGAVGCIPL 132
 Db 4426 HSGFVFSKDKCVPTQCGCKDSG-----TLIPAGKNWITTCGSRCTCTGGLVQCHD 4478
 QY 133 L-CPQLSLPNLGCPNPLVKVTGQC-----CE-----EWVCDSDSIKDPME 173
 Db 4479 FQCPSGAEQDIDGNSNCVEITVQCPAHSYKCLPPCPCSDPDGHCETSPSPST 4538
 QY 174 DDGLGKELGFDASEVELTRNELLIAVGKGRSLKRLPVFGMEPRILYNPLQGGKICVQT 233
 Db 4539 CEEGCV-----CEPDVY--LSNDKCVPS- 4559
 QY 234 TSWSQSKTCGTGISTRTVNDNPECLVKETRICVPRC-----GQPVYS--SLKKG 283
 Db 4560 ---SECGCKDAHGV-----LIPESKTVWSRGCTKNCCTCKGTVQCHDFSCPTG 4604
 QY 284 KCKSKTKKSP-----PVRFTYAGCLSVKRYKRYKCGSCVD--GRC--CTPQLTRT 330
 Db 4605 SRCLDNNEGNSCVTVALKCPAHSLYTNCL-----PCLPSCSDPEGLCGTSEVPST 4658
 QY 331 VQRFRCEDGETFSKNVMMIQ-SCXC 355
 Db 4659 CREGCICSGYVYLHKRCKMLRIHCD 4684

RESULT 12
 A39804
 thrombospondin precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A39804
 R:Lawler, J.; Duquette, M.; Ferro, P.
 J. Biol. Chem. 266, 8039-8043, 1991
 A:Title: Cloning and sequencing of chicken thrombospondin.
 A:Reference number: A39804; MUID:91217026; PMID:2022631
 A:Accession: A39804
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1178 <LAW>
 A:Cross-references: GB:M60853; NID:G212763; PIDN:AAA51437.1; PID:G212764
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
 F:325-383/Domain: von Willebrand factor type C repeat homology <WNC>
 F:386-437/Domain: thrombospondin type 1 repeat homology <THR1>
 F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>

F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F:658-697/Domain: EGF homology <EGF>

Query Match 7.1%; Score 150.5; DB 1; Length 1178;
Best Local Similarity 23.0%; Pred. No. 0.0025;
Matches 65; Conservative 33; Mismatches 102; Indels 83; Gaps 13;

QY 100 CEYNSRIYQNGSFQPNCHOCTIDGAVGCIPLCPQBELSLPNLGCNPNRLVKKVYQCCCE 159
DB 326 CWDQGVFADSESWIVDSCTKCDQSKIVC-----HQITCPVSCADPSFIE--GECPP 378
QY 160 EWVCDBDSIKDPMEDQDGLGKELGDFASEVELTRNNELIAGVGRSLKRLPVFGMEPRI 219
DB 379 --VC---SHSDSEE-----GNSPMSDWTKCVTSGSTQMGRCSDVTRACTG--PHI 426
QY 220 LYNPLOGQKC--IVQT-----TWSQSKTCGIGISTRTVNDNPECLV----- 261
DB 427 QTRMCSFKKCDHRIQDGGHWSHSPSSCVTCGVGNITRILCNPSIPQMGKNVCVNG 486
QY 262 KETRICEVKPC-----IDGLSNPCPPGAECSNYPDGSWSGCP 582
DB 487 RETEKCEKAPCVNGGWPSPWASACTVTCGGIRBSRLCNPSPEPQYGGKPCVGDTHQH 546
QY 284 KCSKTKKSPVRFYACLSVKYRKYPCGSCVDGR--CTTP 325
DB 547 DMCKN-RDCP-----IDGLSNPCPPGAECSNYPDGSWSGCP 582

RESULT 13

A42112
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Oct-1997
C:Accession: A42112
R:Xu, G.; Huan, L.J.; Khatri, I.A.; Wang, D.; Bennick, A.; Fahim, R.E.; Forstner, G.G.;
J. Biol. Chem. 267, 5401-5407, 1992
A:Title: CDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide.
A:Reference number: A42112; MUID:192184794; PMID:1371999

A:Status: preliminary
A:Molecule type: nucleic acid; protein
A:Residues: 1-837 <XU1>
A:Experimental source: intestine
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI Backbone (NCBI:87420)
C:Superfamily: von Willebrand factor type C repeat homology
F:582-650/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.1%; Score 150; DB 2; Length 837;
Best Local Similarity 19.2%; Pred. No. 0.0019;
Matches 88; Conservative 40; Mismatches 138; Indels 192; Gaps 24;

QY 25 TCPA-----CKVC-----AKQLNEDCSKTQPCDH 72
DB 419 TCPAHRQYACGSPSEPTQSSSPKNSTLLVEGCFCEGTTKFAPGYDVCVKICGCVGPD 478

QY 53 -----CKVC-----AKQLNEDCSKTQPCDH 72
DB 479 NVPRFEGHFEDCKDCVLEGGGIVCQPKKARGNLTTCEDGTLYVLEADPDDKCN 538
QY 73 TXGLECN-----FGASSTALKGICRAQSEGRP---CEYNSRIYQNGSFQFP 115
DB 539 TTSCKDPRKCAERPSCLLGFVKSHEVPGKCCPVSCVPKGVCHENAEYQPGSPVYS 598
QY 116 NQCHOCTCID-----GAVGC--IPLCPQBELSLPNLGCNPN-RLVKVYQCCCEWVCD 164
DB 599 NKQCQCVCTSDMNSTQLNVISCTHVP-C-----NISCSSGSELVPEVPGECCKCQQT 650
QY 165 EDSIKDPMED---QDGLGKELG-----FDASEVELTRNNELIAGVGRSLKRLPVFGM 215
DB 651 HCTIKPEQYIILKPEIQKNPNDRCTFFSCMKI-----NNQLIS-----SVSNITCPDF 701
QY 216 EPRILYNPLOGQKICVQTTSW--SQSKTCGIGISTRTVNDNPECLVKEICEVPCG 273

DB 702 DP-----SDCVPGSITYPNPGCKTC-----IHPN-----NTVPCS 733
QY 274 Q-PVYSSLKKGKSKTKSPVPRTYAGCLSVKKYRKYC-GSC-----VD 319
DB 734 AIPVKEI-----SYNGC--AKNISMNFCAGSGCTFAMYSQAQDLID 773
QY 320 G--RCCPTQLTRTVKVRFRCEDEGTFSKNMMIQCKC 355
DB 774 HGCSCCREERTSRVMSLDCPDGSKLSHSVTHIESCLC 811

RESULT 14

A53612
laminin B1 chain precursor - human
A:Alternate names: Kallinin B1 chain; nicein B1 chain
C:Species: Homo sapiens (man)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: A53612
R:Gercke, D.R.; Wagan, D.W.; Champliand, M.F.; Burgeson, R.E.
J. Biol. Chem. 269, 11073-11080, 1994
A:Title: The complete primary structure for a novel laminin chain, the laminin B1 chain
A:Reference number: A53612; MUID:94209274; PMID:7512558
A:Accession: A53612
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1170 <GER>
A:Cross-references: GB:L25541; NID:G510702; PIDN:AAA61834.1; PID:G510703
A:Note: authors translated the codon CGA for residue 124 as Gln, GAG for residue 439 as
C:Superfamily: laminin-type EGF-like homology
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-1170/Product: laminin B1 chain #status predicted <MAT>
F:250-312/Domain: laminin-type EGF-like homology <LEG1>
F:378-427/Domain: laminin-type EGF-like homology <LEG>
F:430-476/Domain: laminin-type EGF-like homology <EG11>
F:532-576/Domain: laminin-type EGF-like homology <LEG2>

Query Match 7.0%; Score 147.5; DB 2; Length 1170;
Best Local Similarity 20.9%; Pred. No. 0.0041;
Matches 84; Conservative 31; Mismatches 101; Indels 185; Gaps 22;

QY 29 AHCPLKAPKAPGVGL-----VRDGGC-----CKVCAKQLN-----ED 63
DB 249 SCFHGHADRCAPKPGASAGSTAVQVHDVVCQHNHTAGNRCERCAFFYNNRWRPAEGD 308
QY 64 CSKTQCDHTKGLCN-----FGASSTALKGI---CRAQSEGRPC---YNS 104
DB 309 AHECQCD-----CNGHSETHCFDPAVFAASQAGVGVCDNCRDHTGKNCERCOLHYF 362
QY 105 RIYQNGSFQPNCHOCTC-IDGAVGCIPLCPQBELSLPNLGCNPNRLVKKVYQCCCEWVC 163
DB 363 RNRRPGASIQETC-ITSCEDPDGAVAGAFCDP-----VTGQC---VC 400
QY 164 DEDSIKDPMEDQDGLGKELGDFASEVELTRNNELIAGVGRSLKRLPVFGMEPRIYNP 223
DB 401 KEH----- 403
QY 224 LQCKKIVQTTSSQSKTCGIGISTRTVNDN-----ECRLV-----KETRICBV 269
DB 404 VQGERC-----DLCKPFG-TGLTYANPRCHRCDNCLGSRMPCDEEGRCJC 451
QY 270 RP-----CGQ--PVYSSLKKGKSKTKSPVPRTYAGCLSVKKYRKYPCGSC--- 317
DB 452 LPNVVGPCKDQCAPYHMKLASGGCEPCACDPN-----SLSPQCNGTGGQPCRE 502
QY 318 -VDGRCTTQLTRTVKVRFRCEDEGTFSKNMMIQCKCNY 357
DB 503 GFGLMCSAAAIR-----QCPD-RTYGDVATGCRACDCDF 536

RESULT 15

T09059

notch4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002
C:Accession: T09059
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Se
submitted to the ENBL Data Library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: Z16543
A:Accession: T09059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1964 <ROW>
A:Cross-references: EMBL:AF030001; NID:G2564945; PID:G2564947
C:Genetics:
A:Gene: notch4
A:Map position: 17
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67
1679/3; 1729/1; 1761/3
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: receptor; signal transduction
F:514-545/Domain: EGF homology <EGF>

Query Match 6.9%; Score 146; DB 2; Length 1964;
Best Local Similarity 23.4%; Pred. NO. 0.0086;
Matches 94; Conservative 28; Mismatches 141; Indels 138; Gaps 24;

QY 28 AACHCPLKAPK--CAPGVGLVDRDGGCKKCAKOLNEDCSKT-----QP-----CD 71
DB 523 AACHLLNGFQCLLPFTGAR-----CEKMDK--CSSTPCANGRCRDQGAIFYCE 573
QY 72 HTKGLK-----CNFGASSTALKG-----ICRAQSEGRPCFY-----NSRIYQN 109
DB 574 CLPFGEPHCEKEVDECLSDPCVAGSCLDLPGAFFCLCRPGFTGQLCEVPLCTPNNMQP 633
QY 110 GESFQPNCOHQ-----CTCIDGAVGCIPL-----CPQELSLNLCNPNRLVKVT 154
DB 634 GQ-----QCQGEHRAFLCPDGPSPGVFAEDNCPCHGHGHCORSICVDEGWTGPECETEL 689
QY 155 GQCEEWVCDSDSIKDPMEDQDGLIGELGFDASEVELTRNNELIYVKGKSLKRLPVFG 214
DB 690 GCCTST-PCAAGGTCHPQPS-----GYNCTCPAGVMGLTCSSEVTACHSGPC----- 735
QY 215 MEPRILYNPLQCKICIVOTTSWSQCKTCGTGISTRTVNDNPECLVKETPRICE--VRPC 272
DB 736 -----LNGGSCSIRPEGYSTCLFSHTG-----RHCQTAYDHC 768
QY 273 GQPVYSSLLKKGKSKTKKSPVPVFTYAGCLSVKVKYRKYC-----GSCVDGSC----- 322
DB 769 ---VSASCLNGTCTVN-----KPGTFF---CLCATGFOGLHCEBKTNPSCADSPCRNKAT 817
QY 323 C--TPQLTRTVQMRPRCEDGETFSKNVMWIOSCKNVCNCPH 361
DB 818 CQDTPRGARCL-----CSPGYTGSSCQTLIDLC-ARKPCPH 852

Search completed: April 22, 2004, 17:59:22

Job time : 17.5434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:52:34 ; Search time 10.5276 Seconds
(without alignments)
1884.444 Million cell updates/sec

Title: US-09-495-448A-4

Perfect score: 2116

Sequence: 1 MSRIARALALVILLHLTR.....ANEAAFPYRLNDIHKFRD 381

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2106	99.5	381	1	O06222 homo sapien
2	1929	91.2	379	1	P18406 mus musculu
3	1929	91.2	379	1	O9es72 rattus norv
4	1695	80.1	375	1	P19336 gallus gall
5	955	45.1	347	1	O9xle9 rattus norv
6	952.5	45.0	348	1	P29268 mus musculu
7	947	44.8	349	1	P29279 homo sapien
8	933	44.1	349	1	O18739 bos taurus
9	906	42.8	349	1	O19113 sus scrofa
10	851.5	40.2	351	1	P28686 gallus gall
11	848	40.1	353	1	P42642 coturnix co
12	827.5	39.1	357	1	P48745 homo sapien
13	825.5	38.0	343	1	P51609 xenopus lae
14	822	38.8	354	1	Q84299 mus musculu
15	818.5	38.7	351	1	O9qzq5 rattus norv
16	767	36.2	367	1	O95388 homo sapien
17	758	35.8	367	1	O54775 mus musculu
18	758	35.8	367	1	O99p90 rattus norv
19	613	29.0	354	1	O55389 homo sapien
20	525.5	24.7	250	1	O76076 homo sapien
21	522.5	24.7	251	1	Q92094 mus musculu
22	490	23.2	250	1	O9jhc6 rattus norv
23	157	7.4	1700	1	Q03376 chironomus
24	153.5	7.3	2282	1	P57999 oryctolagus
25	152	7.2	3110	1	P24043 homo sapien
26	152	7.2	5376	1	O98799 mus musculu
27	150.5	7.1	1178	1	P35440 gallus gall
28	150	7.1	837	1	P98089 rattus norv
29	147.5	7.0	456	1	Q90zds gallus gall
30	146	6.9	1529	1	O94813 homo sapien
31	146	6.9	1964	1	P1695 mus musculu
32	144.5	6.8	1172	1	Q13751 homo sapien
33	143	6.8	1173	1	P35448 xenopus lae

ALIGNMENTS

RESULT 1

CYR6 HUMAN	141	6.7	480	1	HRA1_MOUSE	Q9r118 mus musculu
AC	006222	O149344	O43775	O9BZL7	PRT; 381 AA.	P98088 homo sapien
DT	15-JUL-1998	(Rel. 36, Created)				Q02817 mus sapien
DT	15-JUL-1998	(Rel. 36, Last sequence update)				Q920c1 mus musculu
DT	10-OCT-2003	(Rel. 42, Last annotation update)				Q28178 bos taurus
DE	CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)					Q9y6n6 homo sapien
DE	(Insulin-like growth factor-binding protein 10) (GIGI protein).					P46530 brachydanio
GN	CYR61 OR IGFSP10 OR CCN1 OR GIGI.					Q28833 sus scrofa
OS	Homo sapiens (Human).					P07207 drosophila
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					P83110 homo sapien
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					Q96956 homo sapien
OX	NCBI_TaxID=9606;					
RN	[1]					
RN	SEQUENCE FROM N.A.					
RA	Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;					
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.					
RN	[2]					
RN	SEQUENCE FROM N.A.					
RX	MEDLINE=97280750; PubMed=9135077;					
RA	Jay P., Berge-Leffranc J.L., Marsollier C., Mejean C., Taviaux S.,					
RA	Berta P.;					
RT	"The human growth factor-inducible immediate early gene, CYR61, maps					
RL	to chromosome 1p.";					
RN	[3]					
RN	Oncogene 14:1753-1757(1997).					
RN	SEQUENCE FROM N.A.					
RX	MEDLINE=98197344; PubMed=9536281;					
RA	Martinez C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;					
RT	"Chromosomal mapping and expression of the human cyr61 gene in tumour					
RL	cells from the nervous system.";					
RN	[4]					
RN	Mol. Pathol. 50:310-316(1997).					
RN	SEQUENCE FROM N.A.					
RC	TISUE=Placenta;					
RA	Kolesnikova T.V., Lau L.F.;					
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.					
RN	[5]					
RN	SEQUENCE FROM N.A.					
RA	Bi A.B., Yu L.;					
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.					
RN	[6]					
RN	SEQUENCE FROM N.A.					
RA	Schuetze N., Lechner A., Groll C., Koshrie J., Jakob F.;					
RT	"Regulation of hCYR61 by vitamin D, serum and cytokines in fetal human					
RL	osteoblasts.";					
RN	[7]					
RN	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.					
RN	SEQUENCE FROM N.A.					
RA	Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;					
RT	"Organization and expression of the CYR61 gene in normal human					
RL	fibroblasts.";					
RN	[8]					
RN	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.					
RN	SEQUENCE FROM N.A.					

QY 1 MSSRIARALALVWTLHLTRIALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALALVWTLHLTRIALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALAGICRAOSEGRPCVNSRIYONGESFOPNCHQ 120
 Db 61 NEDCSKTQPCDHTKGLCNFGASSTALAGICRAOSEGRPCVNSRIYONGESFOPNCHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLNPGNPRLVKVTGQCCEEWVCDSEIKDPMEDQDGLLG 180
 Db 121 CTCIDGAVGCIPLCPQELSLNPGNPRLVKVTGQCCEEWVCDSEIKDPMEDQDGLLG 180
 QY 181 KELGPDASEVELTRNNEILVANGKESLAPLVFGNEPRILNPLGOKKCIQVTTNSQCS 240
 Db 181 KELGPDASEVELTRNNEILVANGKESLAPLVFGNEPRILNPLGOKKCIQVTTNSQCS 240
 QY 241 KTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFYTY 300
 Db 241 KTCGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKTKKSPPEVFYTY 300
 QY 301 AGCLSVKKYRKYCGSCVDGRCCTPOLTRTVKMRFRCEGDTFSKNVMTIQSCKNYCNCP 360
 Db 301 AGCLSVKKYRKYCGSCVDGRCCTPOLTRTVKMRFRCEGDTFSKNVMTIQSCKNYCNCP 360
 QY 361 HANEAAFFPYRLFNDRHFRD 381
 Db 361 HANEAAFFPYRLFNDRHFRD 381

RESULT 2
 CYR6 MOUSE
 ID CYR6 MOUSE STANDARD; PRT; 379 AA.
 AC P18406;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cyf61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 GN CYR61 OR IGFBP10 OR CCN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/c; TISSUE=Fibroblast;
 RX MEDLINE=90287146; PubMed=2355916;
 RA O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
 RT "Expression of cyf61, a growth factor-inducible immediate-early
 gene.",
 RL Mol. Cell. Biol. 10:3569-3577(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A/J; TISSUE=Embryonic fibroblast;
 RX MEDLINE=91288203; PubMed=2062642;
 RA Latinkic B.V., O'Brien T.P., Lau L.F.;
 RT "Promoter function and structure of the growth factor-inducible
 immediate early gene cyf61.",
 RL Nucleic Acids Res. 19:3261-3267(1991).
 RN [3]
 RP HEPARIN-BINDING DOMAIN.
 RX MEDLINE=20387398; PubMed=10821835;
 RA Chen N., Chen C.C., Lau L.F.;
 RT "Adhesion of human skin fibroblasts to Cyf61 is mediated through
 integrin alpha 6beta 1 and cell surface heparan sulfate
 proteoglycans.",
 RL J. Biol. Chem. 275:24953-24961(2000).
 CC -!- FUNCTION: Promotes cell proliferation, chemotaxis, angiogenesis
 and cell adhesion. Appears to play a role in wound healing by up-
 regulating, in skin fibroblasts, the expression of a number of
 genes involved in angiogenesis, inflammation and matrix remodeling
 including VEGF-A, VEGF-C, MMP3, TIMP1, uPA, PAI-1 and

integrins alpha-3 and alpha-5 (By similarity). Cyf61-mediated gene
 regulation is dependent on heparin-binding (By similarity). Down-
 regulates the expression of alpha-1 and alpha-2 subunits of
 collagen type-1 (By similarity). Promotes cell adhesion and
 adhesive signaling through integrin alpha-6/beta-1, cell migration
 through integrin alpha-1/beta-5 and cell proliferation through
 integrin alpha-v/beta-3 (By similarity).
 CC -!- SUBUNIT: Interaction with integrins is heparin- and cell-type-
 dependent and promotes cell adhesion (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
 IN LUNG.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
 NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -!- INDUCTION: By growth factors.
 CC -!- SIMILARITY: Belongs to the CCN family.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 VWFC domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M32490; AAA37512.1; -;
 CC EMBL; X56790; AAA40109.1; -;
 CC PIR; A35669; A35669.
 CC MGD; MGI:88613; Cyf61.
 CC GO; GO:0001569; Patterning of blood vessels; IMP.
 CC InterPro; IPR006208; Cys_knot.
 CC InterPro; IPR006207; Cys_knot_C.
 CC InterPro; IPR000867; Insl_gro_fac_pr.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR001007; VWFC.
 CC Pfam; PF00007; Cys_knot; 1.
 CC Pfam; PF00219; IGFBP; 1.
 CC Pfam; PF00090; tsp_1; 1.
 CC Pfam; PF00093; vwc; 1.
 CC SMART; SM00041; CT; 1.
 CC SMART; SM00121; IB; 1.
 CC SMART; SM00209; TSP1; 1.
 CC SMART; SM00214; VWC; 1.
 CC PROSITE; PS01185; CTCK_1; 1.
 CC PROSITE; PS01225; CTCK_2; 1.
 CC PROSITE; PS00222; IGF_BINDING; 1.
 CC PROSITE; PS50092; TSP1; 1.
 CC PROSITE; PS01208; VWFC_1; 1.
 CC PROSITE; PS50184; VWFC_2; 1.
 CC Chemotaxis; Cell adhesion; Growth factor binding; Heparin-binding;
 KW SIGNAL.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61 PROTEIN.
 FT DOMAIN 98 164 VWFC.
 FT DOMAIN 226 271 TSP TYPE-1.
 FT DOMAIN 277 313 HEPARIN-BINDING.
 FT DOMAIN 284 358 CTCK.
 FT DISULFID 284 321 BY SIMILARITY.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MW; FAGB5014B56A8EE9 CRC64;

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred No 1,3e-141;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDRGCGCKVCAKQL 60
Db 1 MSSSTFTLAVAVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDRGCGCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSFALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSFALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPGLGNPRLVKTGQCCCEWVCDSDSKDPMEDDGLLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPGLGNPRLVKTGQCCCEWVCDSDSKDPMEDDGLLG 180
QY 181 KELGFDASEVELTRNNELIAGKGRSLKPLVFGMEPRILYNPL--QGQKCIQVTTWSQ 238
Db 179 --LGLDASEVELTRNNELIAGKGRSLKPLVFGMEPRILYNPL--QGQKCIQVTTWSQ 236
QY 239 CSKTCGTGISTRTVNDNPECLVKEITRICEVRPCQPVYSSLKKGKCKTKKSPVPVF 298
Db 237 CSKTCGTGISTRTVNDNPECLVKEITRICEVRPCQPVYSSLKKGKCKTKKSPVPVF 296
QY 299 TYAGCLSVKKYRPKYCGSCVDGRCTPQTRTVKMRFCEDGETFSKNVMMIQSKCKNYN 358
Db 297 TYAGCSSVKKYRPKYCGSCVDGRCTPQTRTVKMRFCEDGETFSKNVMMIQSKCKNYN 356
QY 359 CPHANEAAPFFYRLFNHDKFRD 381
Db 357 CPHNEAGFRLYSLFNHDKFRD 379
RESULT 3
CYR6_RAT STANDARD; PRT; 379 AA.
AC Q9ES72;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (insulin-like growth factor-binding protein 10).
GN CYR61 OR IGFBP10 OR CCNI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=20435857; PubMed=10852911;
RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
RA Nitsch R.M.;
RT "Muscarinic acetylcholine receptors induce the expression of the
RT immediate early growth regulatory gene CIR61.";
RL J. Biol. Chem. 275:28929-28936(2000).
CC -!- FUNCTION: Promotes cell proliferation, chemotaxis, angiogenesis
CC and cell adhesion. Appears to play a role in wound healing by up-
CC regulating, in skin fibroblasts, the expression of a number of
CC genes involved in angiogenesis, inflammation and matrix remodeling
CC including VEGF-A, VEGF-C, MMP1, MMP3, TIMP1, uPA, PAI-1 and
CC integrins alpha-3 and alpha-5 (By similarity). Cyrel-mediated gene
CC regulation is dependent on heparin-binding (By similarity). Down-
CC regulates the expression of alpha-1 and alpha-2 subunits of
CC collagen type-1 (By similarity). Promotes cell adhesion and
CC adhesive signaling through integrin alpha-6/beta-1, cell migration
CC through integrin alpha-1/beta-5 and cell proliferation through
CC integrin alpha-v/beta-3 (By similarity).
CC -!- SUBUNIT: Interaction with integrins is heparin- and cell-type-
CC dependent and promotes cell adhesion (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the CCN family.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 WFPC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF218568; AAG14964.1; -.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insi_gro_fac_pr.
DR InterPro; IPR000884; TSP1_gro_fac.
DR InterPro; IPR001007; WVF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; WVF; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; WVF; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; WVF_1; 1.
DR PROSITE; PS01184; WVF_2; 1.
KW Chemotaxis; Cell adhesion; Growth factor binding; Heparin-binding;
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 379 CYR61 PROTEIN.
FT DOMAIN 26 97 IGFBP.
FT DOMAIN 98 164 WVF.
FT DOMAIN 226 271 TSP TYPE-1.
FT DOMAIN 277 313 HEPARIN-BINDING (BY SIMILARITY).
FT DOMAIN 284 358 CTCK.
FT DISULFID 284 321 BY SIMILARITY.
FT DISULFID 301 335 BY SIMILARITY.
FT DISULFID 312 351 BY SIMILARITY.
FT DISULFID 315 353 BY SIMILARITY.
FT DISULFID 320 357 BY SIMILARITY.
SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;
Query Match 91.2%; Score 1929; DB 1; Length 379;
Best Local Similarity 90.9%; Pred. No. 1.3e-141;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDRGCGCKVCAKQL 60
Db 1 MSSSTFTLAVAVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDRGCGCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSFALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSFALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPGLGNPRLVKTGQCCCEWVCDSDSKDPMEDDGLLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPGLGNPRLVKTGQCCCEWVCDSDSKDPMEDDGLLG 180
QY 181 KELGFDASEVELTRNNELIAGKGRSLKPLVFGMEPRILYNPL--QGQKCIQVTTWSQ 238
Db 179 --LGLDASEVELTRNNELIAGKGRSLKPLVFGMEPRILYNPL--QGQKCIQVTTWSQ 236
QY 239 CSKTCGTGISTRTVNDNPECLVKEITRICEVRPCQPVYSSLKKGKCKTKKSPVPVF 298
Db 237 CSKTCGTGISTRTVNDNPECLVKEITRICEVRPCQPVYSSLKKGKCKTKKSPVPVF 296
QY 299 TYAGCLSVKKYRPKYCGSCVDGRCTPQTRTVKMRFCEDGETFSKNVMMIQSKCKNYN 358
Db 297 TYAGCSSVKKYRPKYCGSCVDGRCTPQTRTVKMRFCEDGETFSKNVMMIQSKCKNYN 356
QY 359 CPHANEAAPFFYRLFNHDKFRD 381

357 CPHPNEASFRSLFNDIHKFRD 379

DB 357 CPHPNEASFRSLFNDIHKFRD 379

RESULT 4

CE10_CHICK STANDARD; PRT; 375 AA.

AC P19336;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE CEF-10 protein precursor.

GN CCN1.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauaria; Aves; Neognathae; Galliformes; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89145206; PubMed=2537491;

RA Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;

RT "Identification of a phorbol ester-repressible v-src-inducible gene.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).

CC -!- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.

CC -!- SUBCELLULAR LOCATION: Secreted (Probable).

CC -!- INDUCTION: BY V-SRC.

CC -!- SIMILARITY: Belongs to the CCN family.

CC -!- SIMILARITY: Contains 1 IGFBP domain.

CC -!- SIMILARITY: Contains 1 VWFC domain.

CC -!- SIMILARITY: Contains 1 TSP type-1 domain.

CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J04496; AAA48661.1; -.

PIR; A41428; A41428.

InterPro; IPR006208; Cys_knot.

InterPro; IPR006207; Cys_knot_C.

InterPro; IPR000867; Insl_gro_fac_pr.

InterPro; IPR000884; TSP1.

InterPro; IPR001007; VWFC.

Pfam; PF00007; Cys_knot; 1.

Pfam; PF00219; IGFBP; 1.

Pfam; PF00090; tsp; 1; 1.

Pfam; PF00093; vwc; 1.

SMART; SM00041; CT; 1.

SMART; SM00121; IB; 1.

SMART; SM00209; TSP1; 1.

SMART; SM00214; VWC; 1.

PROSITE; PS01185; CTCK_1; 1.

PROSITE; PS01225; CTCK_2; 1.

PROSITE; PS00222; IGF_BINDING; 1.

PROSITE; PS00092; TSP1; 1.

PROSITE; PS01208; VWFC_1; 1.

PROSITE; PS01184; VWFC_2; 1.

Protein factor binding; Signal.

SIGNAL 1 22

FT CHAIN 23 375 CEF-10 PROTEIN.

FT DOMAIN 98 164 VWFC.

FT DOMAIN 223 268 TSP TYPE-1.

FT DOMAIN 281 355 CTCK.

FT DISULFID 281 318 BY SIMILARITY.

FT DISULFID 298 332 BY SIMILARITY.

FT DISULFID 309 348 BY SIMILARITY.

FT DISULFID 312 350 BY SIMILARITY.

FT DISULFID 317 354 BY SIMILARITY.

SQ SEQUENCE 375 AA; 40651 MW; 95F28553BE35D5AE CRC64;

Query Match 80.1%; Score 1695; DB 1; Length 375;

Best Local Similarity 81.0%; Pred. No. 1.4e-123;

Mismatches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSRTARALAVVTLHLTLAL-STCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAQ 59

DB 1 MGSAGARP-ALAAALLCLARLALGSPCPAVCOCPAAAPQCAPGVGLVDPDCCCKVCAQ 59

QY 60 LNECDCKTOPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYONGSPFPNCOH 119

DB 60 LNECDCKTOPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYONGSPFPNCKH 119

QY 120 QCTCIDGAGVCCPLCPQELSLNLCNPNRLVKKVTCQCEEWVCDSDTKDPWEDQDGLL 179

DB 120 QCTCIDGAGVCCPLCPQELSLNLCNPNRLVKKVTCQCEEWVCDSDTKDPWEDQDGLL 179

QY 180 GKELQFDASEVELTRNNELIAVGRSLKRLPVFGMEP--RILYNPLOGQKQKIVOTTSWS 237

DB 178 SKFGLDASEGELTRNNELIAVKG-GLKMLPVFGSEFQSRAPENP---KCIVOTTSWS 232

QY 238 QCSKTCGTGISTRVNDNPECLVKEITRICEVPCGQPVYSLKKGKCKTKKSPVPVR 297

DB 233 QCSKTCGTGISTRVNDNPDCKLIKETRICEVPCGQPVYSLKKGKCKTKKSPVPVR 292

QY 298 FTYAGCLSVKYPKYGSCVDGRCTPQLTQTVKMRPRCEDGETFSKNVMMIQCKCNY 357

DB 293 FTYAGCSSVKYPKYGSCVDGRCTPQLTQTVKMRPRCEDGETFSKNVMMIQCKCNY 352

QY 358 NCPHANEAPPYRLFNDIHKFRD 381

DB 353 NCPHANEAPPYRLFNDIHKFRD 375

RESULT 5

CTGF_RAT STANDARD; PRT; 347 AA.

ID CTGF_RAT STANDARD; PRT; 347 AA.

AC Q9RIE9; Q9WVS1;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Connective tissue growth factor precursor.

GN CTGF OR CCN2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20145935; PubMed=10679821;

RA Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,

RA Marks S.C. Jr., Owen T.A., Popoff S.N.;

RT "Cloning the full-length cDNA for rat connective tissue growth factor:

RT Implications for skeletal development.";

RL J. Cell. Biochem. 77:103-115(2000).

RN [2]

RP SEQUENCE FROM N.A.

RX Tezuka K., Tamatani T.;

RT "Rattus norvegicus connective tissue growth factor.";

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Major connective tissue mitogen secreted by

CC vascular endothelial cells. Promotes proliferation and

CC differentiation of chondrocytes (By similarity). Mediates heparin-

CC and divalent cation-dependent cell adhesion in many cell types

CC including fibroblasts, myofibroblasts, endothelial and epithelial

CC cells (By similarity). Enhances fibroblast growth factor-induced

CC DNA synthesis (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the CCN family.

CC -!- SIMILARITY: Contains 1 IGFBP domain.

CC -!- SIMILARITY: Contains 1 VWFC domain.

CC -!- SIMILARITY: Contains 1 TSP type-1 domain.

```

CC -I SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF120275; RAD39132.1; --
CC EMBL; AB023068; BA82125.1; --
CC InterPro; IPR006208; Cys knot.
CC InterPro; IPR006207; Cys knot C.
CC InterPro; IPR000867; Cys knot C.
CC InterPro; IPR000864; TSPI.
CC InterPro; IPR001007; WFC C.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00219; IGFBP; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; wvc; 1.
CC SMART; SM00041; CT; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSPI; 1.
CC SMART; SM00214; WVC; 1.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC PROSITE; PS00222; IGF_BINDING; 1.
CC PROSITE; PS50092; TSPI; 1.
CC PROSITE; PS01208; WFC_1; 1.
CC PROSITE; PS0184; WFC_2; 1.
CC Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
CC Signal.
CC SIGNAL 1 24
CC CHAIN 25 347
CC DOMAIN 31 98
CC DOMAIN 99 165
CC DOMAIN 196 241
CC DOMAIN 245 347
CC DOMAIN 254 328
CC DISULFID 254 291
CC DISULFID 271 305
CC DISULFID 282 321
CC DISULFID 285 323
CC DISULFID 290 327
CC CONFLICT 35 35
CC CONFLICT 94 94
CC SEQUENCE 347 AA; 37756 MW; CFEIAL9766B7B16 CRC64;

Query Match 45.1%; Score 955; DB 1; Length 347;
Best Local Similarity 46.2%; Pred. NO. 1.4e-66;
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSSRIARALAVTLHLRLAL-STCPAACHPLE-APKAPGVGLVVDGCGCKVCAK 58
DB 1 MLASVAGPVLAVLLLCRTPATGQDCSAQCAAPRCFAGVSLVDGCGCRVCAK 60
QY 59 QLNEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCVSNRYQNGESFPNCQ 118
DB 61 QLQELCTERDPCDHPHGLFCDGFSPANRKGIVCTAK-DGAPCVFSGVYRSGESFSQCK 119
QY 119 HQCTCIDGAVGCTPLCPQELSLNLCNPRVLRVKTGQCCBWCDEBSIKDPMDDQL 178
DB 120 YQCTCLDGAAGVCPVLCSDVRLPSDPCFPFRVRKPGKCCBWCDEP----- 167
QY 179 LKGLGFDASEVELTNNELIAGKGRSLKRL-PVFGMEPRILNPLQCKQICVOTTSWS 237
DB 168 -----KQRTVVGPALAAVRLDITGPDPTWM-----RANLAVQTTSWS 205
QY 238 QCSKTGCTGISTVTDNPECRVLKETRICEVRPCQGPVYSLKKGKCKTKKSPVPR 297
DB 206 ACSKTGCMGISTVTDNPTFCLEKQSLCWMVRPCADLEENIKGKKCIPTPKIAKPVK 265

```

```

QY 298 FTVAGCLSVKRYPKYCGSCVDRGCTCPOLTRTVKMRFCDEGTESKNNMIQSCKNY 357
DB 266 FEISGCTSVKTYRAKCGVCTDRCCTPHRTTLTPVEFKCPDGEIMKKNWFIKTCACHY 325
QY 358 NCPHANEAPPFY--RLFNDI 376
DB 326 NCPGNDNIFESLYRKMYGDM 346

RESULT 6
CTGF_MOUSE STANDARD; PRT; 348 AA.
AC P29268; Q922U0;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor precursor (FISP-12 protein)
DE (Hypertrophic chondrocyte-specific protein 24).
GN CTGF OR CN2 OR FISP12 OR FISP-12 OR RCS24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=91363290; PubMed=1888698;
RA Rybeck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
RT "Structure, mapping, and expression of fisp-12, a growth factor-
RT inducible gene encoding a secreted cysteine-rich protein.";
RL Cell Growth Differ. 2:225-233(1991).
RN 2
RP SEQUENCE FROM N.A.
RX MEDLINE=9129699; PubMed=2029337;
RA Brunner A., Chinn J., Neubauer M.G., Purchio A.P.;
RT "Identification of a gene family regulated by transforming growth
RT factor-beta.";
RL DNA Cell Biol. 10:293-300(1991).
RN 3
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.C., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 4
RP FUNCTION. AND SUBCELLULAR LOCATION.
RX MEDLINE=9327410; PubMed=9184077;
RA Kireeva M.L., Latinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
RA Abler A.S., Lau L.F.;
RT "Cyr61 and Cyrp12 are both ECM-associated signaling molecules;
RT activities, metabolism, and localization during development.";
RL Exp. Cell Res. 233:63-77(1997).
RN 5
RP FUNCTION.
RX MEDLINE=99182484; PubMed=10082563;
RA Babic A.M., Chen C.C., Lau L.F.;
RT "Fisp12/mouse connective tissue growth factor mediates endothelial

```

cell adhesion and migration through integrin alphavbeta3, promotes endothelial cell survival, and induces angiogenesis in vivo.";
 Mol. Cell. Biol. 19:2958-2966(1999).
 !- FUNCTION: Major connective tissue mitogen secreted by vascular endothelial cells. Promotes proliferation and differentiation of chondrocytes (By similarity). Mediates heparin- and divalent cation-dependent cell adhesion in many cell types including fibroblasts, myofibroblasts, endothelial and epithelial cells (By similarity). Enhances fibroblast growth factor-induced DNA synthesis (By similarity).
 !- SUBUNIT: Monomer (By similarity).
 !- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a soluble form.
 !- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 !- INDUCTION: By growth factors.
 !- SIMILARITY: Belongs to the CCN family.
 !- SIMILARITY: Contains 1 IGFBP domain.
 !- SIMILARITY: Contains 1 WFDC domain.
 !- SIMILARITY: Contains 1 TSP type-1 domain.
 !- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements) or send an email to license@ebi.ac.uk).
 EMBL; M70641; AAA37627.1; -.
 EMBL; M70642; AAA37628.1; -.
 EMBL; M80263; AAA73135.1; -.
 EMBL; BC006783; AAH06783.1; -.
 PIR; A40578; A40578.
 MGI; MGI-95537; Ctgf.
 GO; GO:0005578; C:extracellular matrix; IDA.
 GO; GO:0008201; F:heparin binding; IDA.
 GO; GO:0005178; F:integrin binding; IDA.
 GO; GO:0001525; P:angiogenesis; IDA.
 GO; GO:0016477; P:cell migration; IDA.
 GO; GO:0007160; P:cell-matrix adhesion; IDA.
 GO; GO:0008543; P:EGF receptor signaling pathway; IDA.
 GO; GO:0007223; P:integrin-mediated signaling pathway; IDA.
 GO; GO:0001503; P:ossification; IMP.
 InterPro; IPR006208; Cys knot.
 InterPro; IPR006207; Cys knot C.
 InterPro; IPR000367; Insl_gro_fac_pr.
 InterPro; IPR000884; TSP1_gro_fac_pr.
 InterPro; IPR001007; WFDC.
 Pfam; PF00007; Cys_knot; 1.
 Pfam; PF00219; IGFBP; 1.
 Pfam; PF00090; tsp_1; 1.
 Pfam; PF00093; vwc; 1.
 SMART; SM00041; CT; 1.
 SMART; SM00121; IB; 1.
 SMART; SM00209; TSP1; 1.
 SMART; SM00214; VWC; 1.
 PROSITE; PS01185; CTCK_1; 1.
 PROSITE; PS01225; CTCK_2; 1.
 PROSITE; PS00222; IGFBP; 1.
 PROSITE; PS00222; IGFBP; 1.
 PROSITE; PS00092; TSP1; 1.
 PROSITE; PS01208; WFDC_1; 1.
 PROSITE; PS0184; WFDC_2; 1.
 Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding; Signal.
 SIGNAL. 1 25 POTENTIAL.
 FT CHAIN 26 348 CONNECTIVE TISSUE GROWTH FACTOR.
 FT DOMAIN 32 99 IGFBP.
 FT DOMAIN 100 166 WFDC.
 FT DOMAIN 197 242 TSP TYPE-1.
 FT DOMAIN 246 348 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 255 329 CTCK.

FT DISULFID 255 292 BY SIMILARITY.
 FT DISULFID 272 306 BY SIMILARITY.
 FT DISULFID 283 322 BY SIMILARITY.
 FT DISULFID 286 324 BY SIMILARITY.
 FT DISULFID 291 328 BY SIMILARITY.
 FT CONFLICT 161 161 E -> K (IN REF. 1).
 SQ SEQUENCE 348 AA; 37794 NW; 4D7B6D9089174049 CRC64;
 Query Match 45.0%; Score 952.5; DB 1; Length 348;
 Best Local Similarity 46.6%; Pred. No. 2.1e-66;
 Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;
 QY 1 MSRIARALALVVTLLHL-TRLAL-STCPAAACHCELE-APKCAPGVGLVDRGCGCKVCA 57
 DB 1 MLASVAGPISLALVLLALCTREATGQDCSAQCCAAEAAPHPAGVSLVLDGCGCRVCA 60
 QY 58 KQLNEDCSYQPCDHTKGLNCFGASSYALKIGICRAQSEGRCEYNSRIYQNGESFPQNC 117
 DB 61 KQLGELCTERDFCDPHKGLFCDFGSPANRKGIVCTAK-DGAPCVFGGSVYRSGESFPQSC 119
 QY 118 QHQTCTIDGAVGICPLCPQELSLPNLGCNPRLVKVTGQCEEWVCEDESIKDPMEDQDG 177
 DB 120 KYQCTCLDAGVCPVLCSDVRLPSPDCFPFRVXLPKGCCEWVCEDEP----- 168
 QY 178 LLGKELGFDASEVELTRNNELIANGKRSKLK-LPVFGMEPRILYNPLOGQKCIIVTTSW 236
 DB 169 -----KDRTAGVPALAAVRLDTFGDPDTMM-----RANCLVQTTEM 205
 QY 237 SQCSKTCGTGISTRTVNDNPECLVKETRICSVRPGQGVYSSLLKKGKCKSKSPSPV 296
 DB 206 SACSCKTCGNGISTRTVNDNTFRLKQSLCVRPCEADLENIKKKCIPTPIAKPV 265
 QY 297 RFTYAGLSVKYRPKYCGSVDGRCTPQLTRTVYMRFRCDGETFSQNVMMIOSCKN 356
 DB 266 KFLSGCTSVKTYRAKFCVCTDGRCTPHRTTLLPVEFKCPDGEIMKKNMFIKTCAH 325
 QY 357 YNCPHANEAPFPY--RLFNDI 376
 DB 326 YNCPGNDIFESLYRKMIGDM 347
 RESULT 7
 CTGF_HUMAN STANDARD; PRT; 349 AA.
 ID CTGF_HUMAN STANDARD; PRT; 349 AA.
 AC P29279; Q96QX2.
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Connective tissue growth factor precursor (Hypertrophic chondrocyte-specific protein 24).
 DE specific protein 24).
 GN CTGF OR CCN2 OR HGS24.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Umbilical vein endothelial cells;
 RX MEDLINE-91373462; PubMed-1654338;
 RA Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
 RT "Connective tissue growth factor: a cysteine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CBF-10.";
 RL J. Cell Biol. 114:1285-1294 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Umbilical vein endothelial cells;
 RX MEDLINE-93187114; PubMed-1293144;
 RA Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
 RT "Connective tissue growth factor.";
 RL J. Dermatol. 19:642-643 (1992).
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE=Aorta;
RA MEDLINE=97207446; PubMed=9054739;
RA Omar B.S., Werner A., Garnier J.M., Do D.D., Godoy N., Nauck M.,
RA Marz W., Rupp J., Pech M., Luescher T.P.;
RT "Human connective tissue growth factor is expressed in advanced
RT atherosclerotic lesions.";
RL Circulation 95:831-839(1997).
RN [4]
RN SEQUENCE FROM N.A.
RA Cobley V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RN FUNCTION.
RC TISSUE=Chondrocytes;
RA MEDLINE=20080284; PubMed=10614647;
RA Nakanishi T., Nishida T., Shimo T., Kobayashi K., Kubo T.,
RA Tamatani T., Tezuka K., Takigawa M.;
RT "Effects of CTGF/Hcs24, a product of a hypertrophic chondrocyte-
RT specific gene, on the proliferation and differentiation of
RT chondrocytes in culture.";
RL Endocrinology 141:264-273(2000).
RN [6]
RN HEPARIN-BINDING, AND CELL ADHESION.
RA MEDLINE=22442376; PubMed=12553878;
RA Ball D.K., Rachfal A.W., Kemper S.A., Brigstock D.R.;
RT "The heparin-binding 10 kDa fragment of connective tissue growth
RT factor (CTGF) containing module 4 alone stimulates cell adhesion.";
RL J. Endocrinol. 176:R1-R7(2003).
CC -|- FUNCTION: Major connective tissue mitogen secreted by
CC vascular endothelial cells. Promotes proliferation and
CC differentiation of chondrocytes. Mediates heparin- and divalent
CC cation-dependent cell adhesion in many cell types including
CC fibroblasts, myofibroblasts, endothelial and epithelial cells.
CC Enhances fibroblast growth factor-induced DNA synthesis.
CC -|- SUBUNIT: Monomer.
CC -|- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a
CC soluble form (By similarity).
CC -|- ALTERNATIVE PRODUCTS.
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Long;
CC IsoId=P22279-1; Sequence=Displayed;
CC Name=2; Synonyms=Short;
CC IsoId=P22279-2; Sequence=VSP_002460;
CC Note=No experimental confirmation available;
CC -|- SIMILARITY: Belongs to the CCN family.
CC -|- SIMILARITY: Contains 1 IGFBP domain.
CC -|- SIMILARITY: Contains 1 WFEC domain.
CC -|- SIMILARITY: Contains 1 TSP type-1 domain.
CC -|- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92934; AAA91279.1; -;
DR EMBL; X78947; CAA55544.1; -;
DR EMBL; AL354866; CAC44023.1; -;
DR F01; A40551; A40551.
DR Genbank; HGNC:2500; CTGF.
DR MIM; 121009; -;
DR GO; GO:000578; C:extracellular matrix; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0005520; F:insulin-like growth factor binding; TAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR GO; GO:0009611; P:response to wounding; TAS.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot.
DR InterPro; IPR000867; Insl_gro_fac_pr.

DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; TSP1; 1.
DR Pfam; PF00093; VWF; 1.
DR SMART; SMO0041; CT; 1.
DR SMART; SMO0121; IB; 1.
DR SMART; SMO0209; TSP1; 1.
DR SMART; SMO0214; VWF; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01235; CTCK_2; 1.
DR PROSITE; PS00222; IGFBINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
KW Signal; Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 349
FT DOMAIN 33 100
FT DOMAIN 101 167
FT DOMAIN 198 243
FT DOMAIN 247 349
FT DOMAIN 256 330
FT DISULFID 256 293
FT DISULFID 273 307
FT DISULFID 284 323
FT DISULFID 287 325
FT DISULFID 292 329
FT CARBOHYD 28 28
FT CARBOHYD 225 225
FT VARSPIC 172 198
FT CONFLICT 83 83
FT SEQUENCE 349 AA; 38069 MW; 0ECF8470B357EA95 CRC64;
Query Match 44.8%; Score 947; DB 1; Length 349;
Best Local Similarity 46.6%; Pred. No. 5.6e-56;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;
QY 9 LALVYVLLHLTRLAL-STCPAAACHCPL-APKCAPGVGLVDRGCGCKVCVKALNEDCSK 66
DB 11 VAFVLLALCSRAVAVQNGCGPCRPDPAPRCPAGVSLVLDGCGCVCVKALGELCTE 70
QY 67 TQCDHTKGLGECNFGASSTALKGICRAQSEGRCEYNSRIYNGESFPQNCQHCCTIDG 126
DB 71 RDCDPHKGFLCDFGSPANKIGVCTAK-DGAPCFGTVYRSGESFOSCKYQCTCLDG 129
QY 127 AVGCIPLCPQELSLNLCNPNRLVKVTGOCCEWVDEDSIKDPMEDQDGLLGKELGFD 186
DB 130 AVGCMPLCSMDVRLPSPDCFPFRVRLPGKCCSEWVDEP-----KQD----- 172
QY 187 ASEVELTNNELIATVAGKRLKEL-PVFGMEPRILYNPLQGQKCIQVTTSSQCSKTCTG 245
DB 173 -----TVVGPALAAYLEDTFGDPDTMI-----RANCLVQTIEWSACSKTCTGM 215
QY 246 GISTRVTNDPECLRLVKETRICVPRCGQPVYSSLLKGGCKSKTKSPVRYTACLS 305
DB 216 GISTRVTNDNASCRLEKQSLCMVRPCEADLEENIKKGGCKCIRTPIKISKELSGCTS 275
QY 306 VKYRPKYCGSVNDGRCCTPOLTRTVKMRPRCDGTFTFSKNVVMIOCKNCYNCPHANE 365
DB 276 MKTYRAFKGVCTDGRCTTTRTTTLPVEPKCPDGEWVKQNMFIKTCACHYCPGNDI 335
RESULT 8
CTGF_BOVIN
ID CTGF_BOVIN STANDARD; PRT; 349 AA.

RT "Purification and characterization of novel heparin-binding growth
 RT factors in uterine secretory fluids. Identification as heparin-
 RT regulated Mr 10,000 forms of connective tissue growth factor.";
 RL J. Biol. Chem. 272:20275-20282(1997).
 CC -!- FUNCTION: Major connective tissue mitogen secreted by
 CC vascular endothelial cells. Promotes proliferation and
 CC differentiation of chondrocytes (By similarity). Mediates heparin-
 CC and divalent cation-dependent cell adhesion in many cell types
 CC including fibroblasts, myofibroblasts, endothelial and epithelial
 CC cells (By similarity). Enhances fibroblast growth factor-induced
 CC DNA synthesis (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a
 CC soluble form (By similarity).
 CC -!- SIMILARITY: Belongs to the CCN family.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 WFPC domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U83916; AAC48756.1; -.
 DR InterPro; IPR006208; Cys knot.
 DR InterPro; IPR006207; Cys knot C.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WFC C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00019; IGFBP; 1.
 DR Pfam; PF00093; tsp1; 1.
 DR Pfam; PF00093; WFC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF BINDING; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS01208; WFC 1; 1.
 DR PROSITE; PS0184; WFC_2; 1.
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
 KW Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 349
 FT DOMAIN 33 100
 FT DOMAIN 101 167
 FT DOMAIN 198 243
 FT DOMAIN 247 349
 FT DOMAIN 256 330
 FT DISULFID 256 293
 FT DISULFID 273 307
 FT DISULFID 284 323
 FT DISULFID 287 325
 FT DISULFID 292 329
 SQ SEQUENCE 349 AA; 38007 MW; BB510E2B552D4A0 CRC64;
 Query Match 42.8%; Score 906; DB 1; Length 349;
 Best Local Similarity 44.9%; Pred. No. 8.1e-63;
 Matches 167; Conservative 57; Mismatches 108; Indels 40; Gaps 7;
 QY 10 ALVTLHLTRIAL-STCPAACHPL-EAPKAPGVLRDGGCGCKVCALQNECSKT 67
 DB AFVLLALCSRPSQDSCGQCQAGRRACPAVSLVLDGCGCCRLCAQLGELCTER 71
 QY 68 QPCDHTKGLCNFGASSTALKIGICBAQSEGRPCFETNSRIYQNGSFQPCQHQCICDGA 127

Db 72 DPCDPHKGLFCDGSPANKRIGVCTAK-DGACVFGTVYRSGESFQSSCKYQCTCLDGA 130
 QY 128 VGCIPICPOBELSLPNLGCPNPLVKVTGOCCEWUCDESDIKDPMEDQGLLKGELGPDFA 187
 Db 131 VGCVPICSDVRLPSDPCFFPRVRLPGKCCCEWUCDEP----- 169
 QY 188 SEVELTRNNELIANGKRSKLKEL-PVFGMEPRILYNPLQGGKCIQVOTTSWSOCSKTCGTG 246
 Db 170 -----KDHTVVGPAALAAAYLEDTFGDPDTMW-----RANCLVQTTESACSCKTCGNG 216
 QY 247 ISTRVNDNPECLVKETRICVPRPGQPVYSLKKGKCKSKTKKSPPEVRYTACGLSV 306
 Db 217 ISTRVNDNAPFCKLEKQSLKCMVRPCEADLEENIKGKCKCIRTPIKISKEVPELSGCTSV 276
 QY 307 KKYRPYCGSCVDGRCTPQLTRTVQMRPRCDGETFSKNVMVIOCKNKNCPHANEAA 366
 Db 277 KTYRAFCGVCTDGRCTPHRTITLVPVEKCPDGEWVKSMFMFKTCACHYNCPGDNDIF 336
 QY 367 FPPY--RLFNDI 376
 Db 337 ESLYRMYGDM 348
 RESULT 10
 NOV_CHICK
 ID NOV_CHICK STANDARD; PRT; 351 AA.
 AC P28586;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
 GN NOV OR CN3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brown leghorn;
 RK MEDLINE=92107157; PubMed=1309586;
 RA Joliet V., Martinie C., Dambrine G., Plassiat G., Brisac M.,
 RA Crochet J., Perbal B.;
 RT "Proviral rearrangements and overexpression of a new cellular gene
 RT (nov) in myeloblastosis-associated virus type 1-induced
 RT nephroblastomas".
 RL Mol. Cell. Biol. 12:10-21(1992).
 CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
 CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
 CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -!- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -!- SIMILARITY: Belongs to the CCN family.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 WFPC domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

<hr/>					
CC	EMLL; X59284; CRA41975.1; -.				
DR	PIR; S20078; S20078.				
DR	InterPro; IPR006208; Cys_knot.				
DR	InterPro; IPR006207; Cys_knot_C.				
DR	InterPro; IPR000867; Insl_gro_fac_pr.				
DR	InterPro; IPR000884; TSP1_gro_fac.				
DR	InterPro; IPR001007; VWF_C.				
DR	Pfam; PF00007; Cys_knot_1.				
DR	Pfam; PF00219; IGFBP_1.				
DR	Pfam; PF00090; tsp_1; 1.				
DR	Pfam; PF00093; vwc; 1.				
DR	SMART; SM00041; CT; 1.				
DR	SMART; SMC00121; IB; 1.				
DR	SMART; SMC0209; TSP1; 1.				
DR	SMART; SMC0214; VWC; 1.				
DR	PROSITE; PS01185; CTCX_1; 1.				
DR	PROSITE; PS01225; CTCX_2; 1.				
DR	PROSITE; PS00222; IGF_BINDING; 1.				
DR	PROSITE; PS50092; TSP1; 1.				
DR	PROSITE; PS01208; VWF_C; 1.				
DR	PROSITE; PS50184; VWF_C-2; 1.				
KW	Proto-oncogene; Growth factor; Signal.				
FT	SIGNAL 1 24				
FT	CHAIN 25 351				
FT	DOMAIN 31 103				
FT	DOMAIN 104 170				
FT	DOMAIN 201 246				
FT	DOMAIN 258 332				
FT	DISEULFID 258 295				
FT	DISEULFID 275 309				
FT	DISEULFID 286 325				
FT	DISEULFID 289 327				
FT	DISEULFID 294 331				
FT	CARBOHYD 274 274				
SQ	SEQUENCE 351 AA; 38268 MW; 1ECBFA3058C6797 CRC64;				
<hr/>					
Query Match 40.2%; Score 851.5; DB 1; Length 351;					
Best Local Similarity 44.0%; Pred. No. 1.3e-58;					
Matches 164; Conservative 44; Mismatches 104; Indels 61; Gaps 8;					
<hr/>					
QY	11 LVVTLHLTRLALS-----TCPAAC--HCPLEAPKCAPGVGLVRDGGCGCKCAKQLNEDC 64				
DB	12 LLLLLLLLRLPCVESGREAAACRP CGGRCPAEPFPAPGVPALVDGCGCLVCARQGESC 71				
QY	65 SKTOPCDHTKLEGNCFNGASSTALKGICRAQSEGPCEYNRSRIYONGESFPNCQHQTCTI 124				
DB	72 SPLLPDESGLYCDRGPEDDGGAGICMW-LEGDNCVPDGMLYRNETTFPSCKYQCTCR 130				
QY	125 DGAVGCTPLCPQEUSLNLGCNPRLVKVTQCCEEWWCVDESDSKDPMEDODGLLG---- 180				
DB	131 DQGTGCLPRCNLGLLLPGDCFFPRKIEVPECECKWVC-----DPRDEV--ILGGFAM 182				
QY	181 -----KELGFDASEVELTRNNELIAVGKGRSLKRLPVFGWEPRILYNPLQGQKCIVOTT 234				
DB	183 AAYRQEATLGIIDVSD-----SSANCIQTTT 207				
QY	235 SWSQCSKTGPTGISITRYNDNPCELRVKEIRICEVRCGPQVYSLSKKGKCKSKTSKPS 294				
DB	208 EWSACSXSOGMGFSITRYNRNQCEWKQTLCWMRPCENEEPSD-KKGKKCIQTKKSMK 266				
QY	295 PVRTTYAGCLSVKKYRKPKYCGSVDVGGCTTQLTFTVKMRPERCEDGETFSKNVMIOGSK 354				
DB	267 AVREYKNCITSQYTKPRYICGLNDGRCCCTPHNTKTIQVEFRCPQGGKFLKPKMMLINTCV 326				
QY	355 CNYNCPHANEAAF 367				
DB	327 CHGNCPOSNNAFF 339				
<hr/>					
RESULT 11	NOV_COTUA				
ID - NOV_COTJA	STANDARD; PRT; 353 AA.				

Db 8 SLPVLLLLLLLLLLLLRRSEVNGREAPCPRCGRCRPAEPPRCAPGVPVAVLDDGGCCCLVCAR 67
QY 59 QNEDCSKTOPCHTKGLKCNFGASTALKGICRAQSEGRPCOEYNSRIYQNGSEFPQNCQ 118
Db 68 QRGESCLPLPCDESGLYCDRGPEGGGTGICMV-LEGDNCVFDGMIYRNGETFPQSC 126
QY 119 HOCCTGIDGAVGICPLCPQSLPLNLGCPNPLVKVTGOCCEWVCDSDIKPMEDQDGL 178
Db 127 YQCTCRDGGIGLPCRNILGLLPGDPPFRKIEVPGECCKWVC-----EPRDEV--L 178
QY 179 LG-----KELGFDASEVELTRNNELIAGVGRSLKRLPVGFMGPRILYNPLOGQK 228
Db 179 LGGFAMAAYRQBATLGDVSD-----SSAN 203
QY 229 CIVOTTSWQCKTCTGTSTVTDNDNPECLRVKETRICEVPCGQPVYSSLKKGKCKSK 288
Db 204 CIEQTTEWAGCRSCGMSFTVTRNQCCEWVKQRLCMRPECNBPSPD-KGKKKCIIR 262
QY 289 TKKSPFVRFTYAGCLSVKKRYPKYCGSCVDRGCCTPOLTRTVKRFCEDEGETFSKNVM 348
Db 263 TKKMKAVRFYKNCVTSVQYKPRYCGLCNDGRCCTPHNTKIQVEFRCPQCKFLKXPM 322
QY 349 MTQSCCKNCPHANEAF 367
Db 323 LINTCVCHGNCPOSNNAFF 341

RESULT 12
NOV_HUMAN
ID NOV_HUMAN STANDARD; PRT; 357 AA.
AC P48745; Q96BY5.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed gene protein homolog).
GN NOV OR CN3 OR NOVH OR IGFBP9.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94336229; PubMed=7520150;
RA Martinerie C., Huf V., Joubert I., Badzioch M., Saunders G.F.,
RA Strong L.C., Perbal B.;
RT "Structural analysis of the human nov proto-oncogene and expression in Wilms tumor";
RL Oncogene 9:2729-2732(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC Jiang D., Gou D., Li W.;
RT "Cloning, sequencing and expression of human nov gene";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
Satterth A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
INTERACTION WITH FBLN1.
MEDLINE=91128329; PubMed=9927660;
RA Perbal B., Martinerie C., Sainson R., Werner M., He B., Roizman B.;
"The C-terminal domain of the regulatory protein NOVH is sufficient to promote interaction with fibulin 1C: a clue for a role of NOVH in cell-adhesion signaling";
Proc. Natl. Acad. Sci. U.S.A. 96:869-874(1999).
CC -1- FUNCTION: Immediate-early protein likely to play a role in cell growth regulation (By similarity).
CC -1- SUBUNIT: Interacts with FBLN1.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Increased expression in Wilms tumor of the stromal type.
CC -1- SIMILARITY: Belongs to the CCN family.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFEC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL; X78351; CAA55146.1; -
EMBL; X78352; CAA55146.1; JOINED.
EMBL; X78353; CAA55146.1; JOINED.
EMBL; X78354; CAA55146.1; JOINED.
EMBL; X9584; CAA5403.1; -
EMBL; X9584; CAA5403.1; -
EMBL; AY082381; AAL92490.1; -
EMBL; BC015028; AAL15028.1; -
PIR; I38069; I38069.
Genew; HGNC:7885; NOV.
MIM; 164958; -
InterPro; IPR006208; Cys_knot.
InterPro; IPR006207; Cys_knot_C.
InterPro; IPR000867; Insl_gro_fac_pr.
InterPro; IPR000884; TSP1.
InterPro; IPR001007; WVF_C.
Pfam; PF00007; Cys_knot; 1.
Pfam; PF00219; IGFBP; 1.
Pfam; PF00030; tsp_1; 1.
Pfam; PF00093; wvc; 1.
SMART; SM00041; CT; 1.
SMART; SM00121; IB; 1.
SMART; SM00209; TSP1; 1.
SMART; SM00214; WVC; 1.
PROSITE; PS01185; CTCK_1; 1.
PROSITE; PS01235; CTCK_2; 1.
PROSITE; PS00222; IGF_BINDING; 1.
PROSITE; PS00092; TSP1; 1.
PROSITE; PS01208; WVF_C; 1.
PROSITE; PS0184; WFEC_2; 1.
Growth factor; Signal.
SIGNAL 1 27
FT CHAIN 28 357 NOV PROTEIN HOMOLOG.
FT DOMAIN 35 107 IGFBP.
FT DOMAIN 108 174 WFEC.
FT DOMAIN 205 250 TSP TYPE-1.
FT DOMAIN 264 338 CTCK.
FT DISULFID 284 301 BY SIMILARITY.
FT DISULFID 291 315 BY SIMILARITY.
FT DISULFID 292 331 BY SIMILARITY.

RESULT 14
NOV MOUSE
ID NOV MOUSE STANDARD; PRT; 354 AA.
AC 064239; O8CA67;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
DE gene protein homolog).
GN NOV OR CN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv, and ICR; TISSUE=Brain;
RX MEDLINE=97131523; PubMed=8975721;
RA Snaith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,
RA Perbal B., Schofield P.N., Boulter C.A.;
RT "Genomic structure and chromosomal mapping of the mouse nov gene";
RL Genomics 38:425-428(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RX MEDLINE=96304003; PubMed=8622864;
RX Martinerie C., Chevalier G., Rauscher F.J. III, Perbal B.;
RT "Regulation of nov by Wt1: a potential role for nov in
RT nephrogenesis";
RL Oncogene 12:1479-1492(1996).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryonic head, and Spinal cord;
RX MEDLINE=22354683; PubMed=1246851;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RX Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RX Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RX Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RX Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RX Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RX Dalla E., Dragani T.A., Flecher C.F., Forrest A., Frazer K.S.,
RX Gascerland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RX Grimonard S., Gustinchin S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RX Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RX Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RX Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RX Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Peole G.,
RX Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RX Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RX Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RX Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RX Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RX Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RX Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RX Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RX Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RX Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RX Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RX Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=Czech II; TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,
RX Klausner R.D., Collins F.S., Wagner L., Buetow K.H., Bhat N.K.,
RX Altschul S.F., Jordan B., Zeeberg B., Buetow K.H., Bhat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RX Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.B.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Besak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinaki M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Immediate-early protein likely to play a role in cell
CC growth regulation (By similarity).
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the CCN family.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 WFPC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97863; CAA66457.1; -;
CC EMBL; X9257; CAA70454.1; -;
CC EMBL; X96585; CAA65404.1; -;
CC EMBL; AK039481; BAC30363.1; -;
CC EMBL; AK081944; BAC38378.1; -;
CC EMBL; BC003774; AA03774.1; -;
CC MGD; MGI:109185; Nov.
CC InterPro; IPR006208; Cys_knot.
CC InterPro; IPR006207; Cys_knot_C.
CC InterPro; IPR000867; Insi_gro_fac_pr.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00219; IGFBP; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; wvc_1; 1.
CC SMART; SM00041; CT; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSP1; 1.
CC SMART; SM00214; WVC; 1.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC PROSITE; PS00222; IGF_BINDING; 1.
CC PROSITE; PS00092; TSP1; 1.
CC PROSITE; PS01208; WVC_1; 1.
CC PROSITE; PS0184; WVC_2; 1.
CC Growth factor; Signal.
CC SIGNAL 1 21
CC CHAIN 22 354
CC NOV PROTEIN HOMOLOG.
CC FT CHAIN 22 354
CC FT IGFBP.
CC FT DOMAIN 102 168
CC FT WFPC.
CC FT DOMAIN 202 247
CC FT TSP TYPE-1.
CC FT DOMAIN 261 335
CC FT CTCK.
CC FT DISULFID 261 298
CC FT BY SIMILARITY.
CC FT DISULFID 278 312
CC FT BY SIMILARITY.
CC FT DISULFID 289 328
CC FT BY SIMILARITY.
CC FT DISULFID 292 330
CC FT BY SIMILARITY.
CC FT DISULFID 297 334
CC FT BY SIMILARITY.
CC FT CARBOHYD 91 91
CC FT N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 277 277
CC FT N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 6 6
CC R -> G (IN REF. 3; BAC30363).

```

SQ SEQUENCE 354 AA; 38928 MW; 08E8E8CF67829DE CRC64;
Query Match 38.8%; Score 822; DB 1; Length 354;
Best Local Similarity 44.5%; Pred. No. 2.4e-56;
Matches 162; Conservative 46; Mismatches 120; Indels 36; Gaps 9;

QY 7 RALALVVTLLH-LTALSL--CPAAC--HCPLAPKCAPGVGLVDRDGGCKKCAKOLNE 62
Db 8 RCLCLGFLLLHLLSQVSLASLRCPSRCPPKPSISPTCAPGVRSVLDGSCCPVCARQGE 67

QY 63 DCSKTQPCDHTKGLKCNFGASSTALKGICRAQSEGRPCCEYNSRIYQNGESFPQNCQHT 122
Db 68 SCSEMRPCDQSSGLYCDRSADNNETGICMW-PEGDNCVFDGVIYRNGEKFPNCPQYFCT 126

QY 123 CIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGQCCSEWVCDSDSIKDPMEDDGLLQKE 182
Db 127 CRDGGIGCLPRCQLDVLPLPGDPCAPKVAVPGECEKWTGCG-----SSEKTLTG-G 177

QY 183 LGFDASEVELTRNNELIAGVGRSLKRLPVFGMEPRILYNPLQGGKCIIVQTTSWSQCSKT 242
Db 178 LALPAYRPEATVGVEL-----SDSSI-----NCIEQITTEWSACSKS 213

QY 243 CGTGISTRTVNDNPECLRVKETRICEVRPCQG-PVYSLKKGKCKSKTKKSPVRFYTA 301
Db 214 CGMGLSTRVTNRNLQCEMVKQTRLCWVRPCEPGEATDMKGLKRLTKSLKSLHLOFK 273

QY 302 GCLSVKVRPKYKCGSVDRGCTCTPQTRVKNRPFCEDETFSKNVMVMIQSCCKNVCNP 361
Db 274 NCTSLYTKPRFCGICSDGRCCCTPNTKTIQVEFCQLPGQIILKPKVWVIGTCTCHSNCP 333

QY 362 ANEA 365
Db 337 NNEA 340

RESULT 15
NOV_RAT
ID NOV_RAT STANDARD; PRT; 351 AA.
AC Q90Z05
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (Novh) (Nephroblastoma overexpressed
DE gene protein homolog).
GN NOV OR CCN3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINS=20035752; PubMed=10570975;
RA Liu C., Liu X.J., Crowe P.D., Keiner G.S., Fan J., Barry G., Manu F.,
RA Ling N., De Souza E.B., Maki R.A.;
RT "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
RT that induces protein tyrosine phosphorylation.";
RL Gene 238:471-478(1999).
CC -!- FUNCTION: Can act as a growth factor for some cells and binds to a
CC specific receptor that leads to the phosphorylation of a 221 kDa
CC protein.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the CGN family.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF171936; AAD49371.1; -
CC InterPro; IPR006208; Cys knot.
CC InterPro; IPR006207; Cys knot_C.
CC InterPro; IPR000867; InaI_gro_fac_pr.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR001007; WFC_C.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00219; IGFBP; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; wfc; 1.
CC SMART; SM00041; CT; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSP1; 1.
CC SMART; SM00214; WVC; 1.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC PROSITE; PS00222; IGF_BINDING; 1.
CC PROSITE; PS00092; TSP1; 1.
CC PROSITE; PS01208; WFC_1; 1.
CC PROSITE; PS0184; WFC_2; 1.
CC Growth factor; Signal.
KW SIGNAL
FT CHAIN 1 21
FT CHAIN 22 351
FT DOMAIN 29 101
FT DOMAIN 102 168
FT DOMAIN 199 244
FT DOMAIN 258 332
FT DISULFID 258 325
FT DISULFID 275 309
FT DISULFID 286 325
FT DISULFID 289 327
FT DISULFID 294 331
FT CARBOHYD 91 91
FT CARBOHYD 274 274
FT SEQUENCE 351 AA; 38509 MW; 02619707DB7C1BFB CRC64;
NOV PROTEIN HOMOLOG.
IGFBP.
WFC.
TSP TYPE-1.
CTCK.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
Query Match 38.7%; Score 818.5; DB 1; Length 351;
Best Local Similarity 43.4%; Pred. No. 4.5e-56;
Matches 158; Conservative 49; Mismatches 118; Indels 39; Gaps 8;

QY 7 RALALVVTLLH-LTALSL--CPAAC--HCPLAPKCAPGVGLVDRDGGCKKCAKOLNE 62
Db 8 RCLCLGFLLLHLLNQVSLATLRCPSRCPPKPSISPTCAPGVRSVLDGSCCPVCARQGE 67

QY 63 DCSKTQPCDHTKGLKCNFGASSTALKGICRAQSEGRPCCEYNSRIYQNGESFPQNCQHT 122
Db 68 SCSEMRPCDQSSGLYCDRSADNNETGICMW-PEGDNCVFDGVIYRNGEKFPNCPQYHCT 126

QY 123 CIDGAVGCIPLCPQELSPLNLCNPNRLVKVTGQCCSEWVCDSDSIKDPMEDDGLLQKE 182
Db 127 CRDGGIGCVRCQLDVLPLPGDPCAPKVAVPGECEKWTGCG-----SSEKTLTG-G 177

QY 183 LGFDASEVELTRNNELIAGVGRSLKRLPVFGMEPRILYNPLQGGKCIIVQTTSWSQCSKT 242
Db 178 LALPAYRPEATVGVEL-----SDSSI-----NCIEQITTEWSACSKS 213

QY 243 CGTGISTRTVNDNPECLRVKETRICEVRPCQG-PVYSLKKGKCKSKTKKSPVRFYTA 301
Db 214 CGMGLSTRVTNRNLQCEMVKQTRLCWVRPCEPGEATDMKGLKRLTKSLKSLHLOFK 273

QY 302 GCLSVKVRPKYKCGSVDRGCTCTPQTRVKNRPFCEDETFSKNVMVMIQSCCKNVCNP 361
Db 274 NCTSLYTKPRFCGICSDGRCCCTPNTKTIQVEFCQLPGQIILKPKVWVIGTCTCHSNCP 333

QY 362 ANEA 365
Db 334 NNEA 337
```

Search completed: April 22, 2004, 17:57:05
Job time : 11.5276 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:53:04 ; Search time 38.6013 Seconds
(without alignments)
3114.206 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALAVVILLHLTR.....ANEAAPFYRLFNHDKFRD 381

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1928	91.1	379	11 Q9WTM9	Q9WTM9 rattus norv
2	1663.5	78.6	375	13 Q98TX5	Q98TX5 xenopus lae
3	1587	75.0	334	4 Q9UID7	Q9UID7 homo sapien
4	938	44.3	343	13 Q42607	Q42607 xenopus lae
5	930	44.0	344	13 Q98T08	Q98T08 gallus gall
6	912.5	43.1	347	13 Q9PT80	Q9PT80 notophthalm
7	878	41.5	349	6 Q97765	Q97765 sus scrofa
8	626.5	29.6	331	4 Q95958	Q95958 homo sapien
9	485.5	22.9	176	13 Q9PS86	Q9PS86 gallus gall
10	373.5	17.7	128	11 Q9R2C0	Q9R2C0 rattus norv
11	346	16.4	125	6 Q862T0	Q862T0 bos taurus
12	341	16.1	113	11 Q9Z164	Q9Z164 rattus norv
13	337	15.9	100	4 Q9UDL6	Q9UDL6 homo sapien
14	336	15.9	119	11 Q91V29	Q91V29 mus musculu
15	336	15.9	119	11 Q920W6	Q920W6 mus spicile
16	302	14.3	374	5 Q9VVK0	Q9VVK0 drosophila

ALIGNMENTS

RESULT 1

ID	Q9WTM9	PRELIMINARY;	PRT;	379 AA.
AC	Q9WTM9;			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	CYR61 precursor.			
GN	CYR61.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=izm; TISSUE=Aorta;			
RA	Unoki H., Yonekura H., Furukawa K., Yamamoto H.;			
RT	"Rat Cyr61 mRNA.";			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB015877; BAA78339.1; -			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0005520; F:insulin-like growth factor binding; IEA.			
DR	GO; GO:0001558; P:regulation of cell growth; IEA.			
DR	InterPro; IPR006208; Cys_knot.			
DR	InterPro; IPR006207; Cys_knot_C.			
DR	InterPro; IPR000867; Insl_gro_fac_pr.			
DR	InterPro; IPR000884; TSP1.			
DR	InterPro; IPR001007; VWF_C.			
DR	Pfam; PF00007; Cys_knot; 1.			
DR	Pfam; PF00219; IGFEP; 1.			
DR	Pfam; PF00030; tsp_1; 1.			
DR	Pfam; PF00033; vwc; 1.			
DR	SMART; SM00041; CT; 1.			
DR	SMART; SM00121; IB; 1.			
DR	SMART; SM00209; TSP1; 1.			
DR	SMART; SM00214; VWC; 1.			
DR	PROSITE; PS01185; CTCK 1; 1.			
DR	PROSITE; PS01235; CTCK 2; 1.			
DR	PROSITE; PS00222; IGF_BINDING; 1.			

Q9VVK3 drosophila
Q8WYK7 homo sapien
Q9UDE4 xenopus lae
Q9IBG7 xenopus lae
Q9DF21 scyllorhinu
Q8T3A6 caenorhabdi
Q8T3A7 caenorhabdi
Q9XWD6 caenorhabdi
Q9NZV1 homo sapien
Q86XX4 homo sapien
Q9TVQ2 caenorhabdi
Q94446 chironomus
Q9110 mus musculu
Q8AW5 gallus gall
Q97574 bos taurus
Q8CG65 mus musculu
Q864N4 caenorhabdi
Q21281 caenorhabdi
Q86AK7 dictyosteli
Q80T14 mus musculu
Q92513 mus musculu
Q8VEA6 mus musculu
Q8SPM4 bos taurus
Q90Z43 gallus gall
Q90XG4 gallus gall
Q8CIZ8 mus musculu
Q9CVG8 mus musculu
Q8XQ1 mus musculu
Q80Z17 mus musculu

DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsg_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00209; TSPI; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00092; TSPI_1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS01208; VWC_2; 1.
 DR PROSITE; PS0184; VWC_2; 1.
 SQ SEQUENCE 334 AA; 37246 MW; 9188987A7352E948 CRC64;

Query Match 75.0%; Score 1587; DB 4; Length 334;
 Best Local Similarity 85.8%; Pred. No. 9.2e-145;
 Matches 295; Conservative 5; Mismatches 22; Indels 22; Gaps 4;

QY 49 GCGCKVCAKQINEDCSKTPQCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQ 108
 DB 2 GCGTHPNCILHIGHTASPTSYKHHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQ 61
 QY 109 NGESFQPCNCHOCCTIDGAVGCIPLCQELSLPILGNPRLVKTGCCCEWVCDSDSI 168
 DB 62 NGESFQPCNCHOCCTIDGAVGCIPLCQELSLPILGNPRLVKTGCCCEWVCDSDSI 121
 QY 169 KPMEDQDGLLKGELGFDASEVELTRNNELIAGVGRSLKRLPVFGMEPRILYNPLQGGK 228
 DB 122 KPMEDQDGLLKGELGFDASEVELTRNNELIAGVGRSLKRLPVFGMEPRILYNPLQGGK 181
 QY 229 CIVQTSNSQSKCTGCTGISTRVNDNPECLVKETRICVPRGQGVYSLKGGKCKSK 288
 DB 182 CIVQTSNSQSKCTGCTGISTRVNDNPECLVKETRICVPRGQGVYSLKGGKCKSK 241
 QY 289 TKKSPVPRFTYAGLSVKKYRKYCGSCVDGRCTPQLTRTVKMRPCED-----GE 341
 DB 242 TKKSPVPRFTYAGLSVKKYRKYCGSCVDGRCTPQLTRTVKMRPCED-----GE 295
 QY 342 TFSKN-----VNMISCKNVCNCPHANEAPFPYFLFNDIHKFRD 381
 DB 296 IFQERHDDPVL-----KCNVNCNCPHANEAPFPYFLFNDIHKFRD 334

RESULT 4
 Q42607 PRELIMINARY; PRT; 343 AA.
 ID O42607
 AC O42607
 DT 01-JAN-1998 (TremBLrel. 05, Created)
 DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Connective tissue growth factor xCTGF.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ying Z.; King M.L.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U43524; AAB67639.1; -
 DR EMBL; U43523; AAB67638.1; -
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005520; F:insulin-like growth factor binding; IEA.
 DR GO; GO:001558; F:regulation of cell growth; IEA.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot.C.
 DR InterPro; IPR000887; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSPI.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsg_1; 1.
 DR Pfam; PF00093; vwc; 1.

DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSPI; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00092; TSPI_1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS01208; VWC_2; 1.
 DR PROSITE; PS0184; VWC_2; 1.
 SQ SEQUENCE 343 AA; 37966 MW; 93F221C5DB565A81 CRC64;

Query Match 44.3%; Score 938; DB 13; Length 343;
 Best Local Similarity 47.2%; Pred. No. 4e-82;
 Matches 168; Conservative 51; Mismatches 95; Indels 42; Gaps 6;

QY 26 CPAACHCPLKAPKAPGVLVVDGGCGCKVCAKQINEDCSKTPQCDHTKGLCNFGASST 85
 DB 24 CNBECQCPNKVPVCDPGRVMVQDGGCGCKVCAKQINEDCSKTPQCDHTKGLCNFGASST 83
 QY 86 ALKGICRAQSEGRPCVNSRIYQNGESFQPCNCHOCCTIDGAVGCIPLCPOELSPLNLGC 145
 DB 84 RLIGVCTAR-EGAPCVFGTVYRSGESFQSSCKYQCTCIDGGVGVCPICSMDIRLPSPEC 142
 QY 146 PNPRLVKKVGGCCCEWVCDSDSIKPMEDQDGLLKGELGFDASEVELTRNNELIAGVGR 205
 DB 143 PFPFRVYKLPKCCCEWVCDP-----QERTLVGPA- 172
 QY 206 SLKRLPVFGMEPRILYNP---LOGOKCIVQTSNSQSKCTGCTGISTRVNDNPECLVK 262
 DB 173 ---LPAPRMEE--TYGDDPSLIRANCLVQTVNSACSKTCMGISTRTVNDNEHCRLEK 226
 QY 263 ETRICEVPRGQGVYSLKGGKCKSKTKKSPVPRFTYAGLSVKKYRKYCGSCVDGRCC 322
 DB 227 QSRLCVWRPCEADLEENIKGKKCIKRTPKISKVPKFEFGCTSVKTYRAKFGVCTDGRCC 286
 QY 323 CTPOLTTRTVKMRPCEDGETSFKNVMMISCKNVCNCPHANE--AAPPYFLFNDI 376
 DB 287 CUPHRTATILPVEFKCPDGEVKKNNMFIKTCACHFNCFGDNDIFEMMYRKYMGDM 342

RESULT 5
 Q98TQ8 PRELIMINARY; PRT; 344 AA.
 ID Q98TQ8
 AC Q98TQ8
 DT 01-JUN-2001 (TremBLrel. 17, Created)
 DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Connective tissue growth factor precursor (Connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gygi D.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mukudai Y.; Kubota S.; Takigawa M.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ298335; CAC33438.1; -
 DR EMBL; AF463517; AAL68834.1; -
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005520; F:insulin-like growth factor binding; IEA.
 DR GO; GO:0001558; F:regulation of cell growth; IEA.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSPI.
 DR InterPro; IPR001007; VWF.C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00090; tsg_1; 1.
 DR Pfam; PF00093; vwc; 1.


```

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92096871; PubMed=1756408;
RA Martinie C., Perbal B.;
RT "Expression of a gene encoding a novel potential IGF binding protein
RT in human tissues."
RL C. R. Acad. Sci., II, Sci. Vie 313:345-351(1991).
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005520; P:insulin-like growth factor binding; IEA.
DR GO; GO:0001558; P:regulation of cell growth; IEA.
DR InterPro; IPRO00867; Inal_gro_fac_pr.
DR InterPro; IPRO01007; VWF_C.
DR SMART; SM00121; IB; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS01208; VWF_C_1; 1.
DR PROSITE; PS0184; VWF_C_2; 1.
FT NON_TER 1
FT NON_TER 77
FT NON_TER 154
FT NON_TER 176
FT NON_TER 176
SQ SEQUENCE 176 AA; 18656 MW; 4FA69FFDB79B1C67 CRC64;

Query Match 22.9%; Score 485.5; DB 13; Length 176;
Best Local Similarity 40.2%; Pred. No. 8.9e-39;
Matches 94; Conservative 21; Mismatches 52; Indels 67; Gaps 5;

QY 24 STCPAAC--HCPLEAPKAPGVLYRDGCGCKYCAKQLNEDCSKTPCDHTKLECNFG 81
DB 8 AACPPCGRCFAEPKAPGVAVLDGCGCLVCARQSGCSPLLPDESGLYCDRG 67
QY 82 ASSTALKGICRAQSGRCEVNSRIYONGESFQNCQHCCTCIDGAVGCIPLCPQELSLP 141
DB 68 PEDGGAGIC---EGDNCVDFGMIYRNGEITFQSKYQCTCRDQIGCLPRCNLGLLP 123
QY 142 NLGCNPRLVKVTCGCEWVDEDSIKDPMEDQGLLGKELGFDASEVELTRNNELIAV 201
DB 124 GPDCFPFKIEVPGCECKWVC-----DPRDEV--LLG----- 154
QY 202 GKGRSLKRLPVFGMEPRILYNPLOGKQKIVQTSWSQCSKTCGTGISTRVTNDN 255
DB 155 -----TEWSACSKSCGMGFSTRVTNRN 176

RESULT 10
QYR2C0 PRELIMINARY; PRT; 128 AA.
AC QYR2C0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE "Connective tissue growth factor is produced by hepatic stellate cells
DE and upregulated in rat liver fibrosis."
DE Connective tissue growth factor (Fragment).
GN CTGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX Jia J.D., Sedlaczek N., Bauer M., Wischer D., Boigk G., Cho J.J.,
RA Ruehl M., Riecken E.O., Schuppan D.;
RT "Connective tissue growth factor is produced by hepatic stellate cells
RT and upregulated in rat liver fibrosis."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ236872; CAB41996.1;
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO01007; VWF C.
DR Pfam; PF00090; tsg_1; 1.

```

```

DR Pfam; PF00093; VWC; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF_C_1; 1.
DR PROSITE; PS0184; VWF_C_2; 1.
FT NON_TER 1
FT NON_TER 128
FT NON_TER 128
SQ SEQUENCE 128 AA; 13803 MW; F30142D0C491B7C7 CRC64;

Query Match 17.7%; Score 373.5; DB 11; Length 128;
Best Local Similarity 43.4%; Pred. No. 4e-28; 34; Indels 35; Gaps 3;
Matches 69; Conservative 21; Mismatches 21;

QY 96 EGRPCVNSRIYONGESFQNCQHCCTCIDGAVGCIPLCPQELSLPNLGCNPRLVKVKG 155
DB 4 DGAPCVFGSVYRSGESFQSCYQCTCLDGAAGCVPLCSMDVRLPSPDCPPRVRKLP 63
QY 156 QCCEWVDEDSIKDPMEDQGLLGKELGFDASEVELTRNNELIAVGGKRSKLRL-PVFG 214
DB 64 KCCEWVDEP-----KDRTVGFPALAAAYRLDFTG 94
QY 215 MEPRILYNPLOGKQKIVQTSWSQCSKTCGTGISTRVTN 253
DB 95 PDPTM-----RANCLVQTTENSACSKTCGMGISTRVTN 128

RESULT 11
QYR2C0 PRELIMINARY; PRT; 125 AA.
AC QYR2C0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to connective tissue growth factor precursor (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwata H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaika Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
RT pregnancy using a custom cDNA microarray."
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB098897; HAC56387.1;
DR InterPro; IPRO06208; Cys_knot.
DR InterPro; IPRO06207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot; 1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
FT NON_TER 125
FT NON_TER 125
SQ SEQUENCE 125 AA; 14352 MW; B6C7A27644DBFAE7 CRC64;

Query Match 16.4%; Score 346; DB 6; Length 125;
Best Local Similarity 46.0%; Pred. No. 1.8e-25;
Matches 57; Conservative 26; Mismatches 39; Indels 2; Gaps 1;

QY 255 NPECLVETRICVVRPCQPVYSSLYKGGKCKSKTKSPFVPTVYAGCLSVKKYRKYC 314
DB 1 NAFCRLEKQSLCMVRPCEADLEENIKGGKICIRTPKISKPIKFLSGCTSKMKTTRAKPC 60
QY 315 GSCVDGRCTCTPQTRTVKMRFCEDGETFSKNVMIQSCKNYCNCPHANEAFPPY--RL 372
DB 61 GVTGDRCTCTPRTTTLTPVEFKCPDGEVWKSKSMFIKTCACHYNCGDNDIESLYYRKM 120
QY 373 FNDI 376
DB 121 YGDM 124

```

```

RESULT 12
Q92164 ID Q92164 PRELIMINARY; PRT; 113 AA.
AC Q92164;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-2003 (TRENBLrel. 23, Last annotation update)
DE Connective tissue growth factor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99150373; PubMed=10026205;
RA Murphy M., Godson C., Cannon S., Kato S., Mackenzie H.S., Martin F.,
RA Brady H.R.;
RT "Suppression subtractive hybridization identifies high glucose levels
RT as a stimulus for expression of connective tissue growth factor and
RT other genes in human mesangial cells.";
RL J. Biol. Chem. 274:5830-5834(1999).
DR EMBL; AF079531; AAD02836.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot_1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12767 MW; 23EA69FC0A60635E CRC64;

Query Match 16.1%; Score 341; DB 11; Length 113;
Best Local Similarity 55.1%; Pred. No. 4.8e-25;
Matches 59; Conservative 17; Mismatches 31; Indels 0; Gaps 0;

QY 247 ISTRVNDNPECLVKETRICVPCGQPVYSSLKKGKCKTKKSPPEVFTVAGCLSV 306
DB 1 ISTRVNDNPTCLLEKQSLQWVRPEADLEENIKGKKCIIRTPKIAKPVKFLSGCTSV 60
QY 307 KKYRPKYCGSVDGRCTPQLTRTVKMFRCEDGETSKNVMMIQSC 353
DB 61 KTYRAKFGVCTDGRCCPTHTTLPVEFKPHEIMKQNMFIKTC 107

RESULT 13
Q9UDL6 ID Q9UDL6 PRELIMINARY; PRT; 100 AA.
AC Q9UDL6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Potential IGF binding protein (Fragments).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92096871; PubMed=1756408;
RA Martinerie C., Perbal B.;
RT "Expression of a gene encoding a novel potential IGF binding protein
RT in human tissues.";
RL C. R. Acad. Sci., III, Sci. Vie 313:345-351(1991).
DR InterPro; IPR001007; VWF_C.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF_C.1; 1.
DR PROSITE; PS0184; VWF_C.2; 1.
FT NON_TER 1
FT NON_TER 79
FT NON_TER 100
SQ SEQUENCE 100 AA; 10813 MW; C64EB2B9AD8A3299 CRC64;

```

```

Query Match 15.9%; Score 337; DB 4; Length 100;
Best Local Similarity 40.0%; Pred. No. 1e-24;
Matches 64; Conservative 17; Mismatches 19; Indels 60; Gaps 3;

QY 96 EGRPCVNSRIYONGSFQPCNQHOCTCIDGAVGCIPLCPQELSLNLCENPRLVKVVG 155
DB 1 DGAPCIFGGTVYNSGFSFSSCKYQCTCIDGAVGCMPLCSMDVRLPSPDCPPRRVKLP 60
QY 156 QCCEWVCDSDSIKDPMDQDGLLGKELGFDASEVELTRNELLIAVGKRSKLRLPFGM 215
DB 61 KCCEWVCDSDSIKDPMDQDGLLGKELGFDASEVELTRNELLIAVGKRSKLRLPFGM 215
QY 216 EPRILYNPLQGGKICIVQTTSMQCSKTCGTGISTRTVNDN 255
DB 79 -----TEWSACSKTCGISTRTVNDN 100

RESULT 14
Q91V29 ID Q91V29 PRELIMINARY; PRT; 119 AA.
AC Q91V29;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE FISP-12 protein (Fragment).
OS FISP-12.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Various strains;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=pgn2;
RA Kobayakawa H., Saitou N.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=pgn2;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039094; BAB68618.1; -
DR EMBL; AB039095; BAB68619.1; -
DR EMBL; AB039096; BAB68620.1; -
DR EMBL; AB039097; BAB68621.1; -
DR EMBL; AB039098; BAB68622.1; -
DR EMBL; AB039099; BAB68623.1; -
DR EMBL; AB039100; BAB68624.1; -
DR EMBL; AB039101; BAB68625.1; -
DR EMBL; AB039102; BAB68626.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot_1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
SQ SEQUENCE 119 AA; 13626 MW; C7C8AD253DF331CE CRC64;

Query Match 15.9%; Score 336; DB 11; Length 119;
Best Local Similarity 47.9%; Pred. No. 1.5e-24;
Matches 56; Conservative 23; Mismatches 36; Indels 2; Gaps 1;

QY 262 KETRICVPCQPVYSSLKKGKCKTKKSPPEVFTVAGCLSVKRYKVCSCVDGR 321

```

Db 2 KOSRLCWRPCEADLEENIKGKKICIRTPKIAKPVKFLSGCTSVKTYRAKFCGVCCTDGR 61
QY 322 CCTPQLTRIVYKMFRCEDGETSKNYMMIQSKCNKYNCPHANEAAFPFY--RLFNDI 376
Db 62 CCTPHRTTILPVFEKCPDGEIMKKNMMFIKTCACHYNCPCGNDNIFESLYYRKMYGDM 118

RESULT 15

Q920W6 PRELIMINARY; PRT; 119 AA.
ID Q920W6;
AC Q920W6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FISP-12 protein (Fragment).
GN FISP-12.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZEN;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AS039103; BAB68627.1; -.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot; 1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
SQ SEQUENCE 119 AA; 13626 MW; C7C8AD253DF331CE CRC64;

Query Match 15.9%; Score 336; DB 11; Length 119;
Best Local Similarity 47.9%; Pred. NO. 1.5e-24;
Matches 56; Conservative 23; Mismatches 36; Indels 2; Gaps 1;

QY 262 KETRICVPRCGQPVYSSLLKGGKSKTKKSPFPVFTYAGCLSVKYPKPCGSCVDGR 321
Db 2 KOSRLCWRPCEADLEENIKGKKICIRTPKIAKPVKFLSGCTSVKTYRAKFCGVCCTDGR 61
QY 322 CCTPQLTRIVYKMFRCEDGETSKNYMMIQSKCNKYNCPHANEAAFPFY--RLFNDI 376
Db 62 CCTPHRTTILPVFEKCPDGEIMKKNMMFIKTCACHYNCPCGNDNIFESLYYRKMYGDM 118

Search completed: April 22, 2004, 17:58:35
Job time : 39.6013 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:54:44 ; Search time 16.4566 Seconds
(without alignments)
1188.962 Million cell updates/sec

Title: US-09-495-448A-2
Perfect score: 2103
Sequence: 1 MSSSTFRTLAVAVTLHLTR.....PNEASFRLYSFLNDIHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	1	US-08-468-847B-11
2	2103	100.0	379	4	US-09-142-569-2
3	1938	92.2	381	4	US-09-348-815-2
4	1929	91.7	381	4	US-09-142-569-4
5	1819.5	86.5	374	1	US-08-468-847B-12
6	1812.5	86.2	375	2	US-08-459-101A-2
7	1646.5	78.3	375	1	US-08-468-847B-13
8	957.5	45.6	347	4	US-09-582-337-2
9	957.5	45.5	348	4	US-09-292-036-3
10	953.5	45.3	348	1	US-08-468-847B-15
11	953.5	45.3	348	4	US-09-142-569-6
12	950	45.2	349	1	US-08-167-628-2
13	950	45.2	349	1	US-08-386-680-2
14	950	45.2	349	1	US-08-459-717-2
15	950	45.2	349	1	US-08-712-302-2
16	950	45.2	349	2	US-08-880-031-2
17	950	45.2	349	3	US-09-054-368-2
18	950	45.2	349	3	US-09-097-179-2
19	950	45.2	349	3	US-09-054-274-2
20	950	45.2	349	3	US-09-080-715-2
21	950	45.2	349	3	US-09-056-704-2
22	950	45.2	349	4	US-09-292-036-4
23	950	45.2	349	4	US-09-253-316-26
24	950	45.2	349	4	US-09-142-569-8
25	950	45.2	349	4	US-09-461-688-2
26	950	45.2	349	5	PCT-US96-08140-2
27	944	44.9	347	4	US-09-187-478-2

Sequence 2, Appl
Sequence 14, Appl
Sequence 16, Appl
Sequence 17, Appl
Sequence 25, Appl
Sequence 3, Appl
Sequence 6, Appl
Sequence 8, Appl
Sequence 5, Appl
Sequence 21, Appl
Sequence 7, Appl
Sequence 22, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 36, Appl
Sequence 37, Appl
Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-468-847B-11
; Sequence 11, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-11

Query Match 100.0%; Score 2103; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5e-170;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSSTFRTLAVAVTLHLTRLALSTCPAACHPCAPGVGLVRDCCGCKVKAKQL 60
|||||

Db 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPCLEAPKCAPGVGLVRDGGCGCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCCEVNSRIYQNGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCCEVNSRIYQNGESFQPNCKHQ 120
Qy 121 CTICIDGAVGCIPLCPQELSPLNLCNPNRPLVKVSGGCCBEEWDCDESIKSLDDDDLLG 180
Db 121 CTICIDGAVGCIPLCPQELSPLNLCNPNRPLVKVSGGCCBEEWDCDESIKSLDDDDLLG 180
Qy 181 LDASEVELTNNELIATGKSSLKRLPVGTEPRVLFNPLHAHQKCIQVTTTWSQCCKS 240
Db 181 LDASEVELTNNELIATGKSSLKRLPVGTEPRVLFNPLHAHQKCIQVTTTWSQCCKS 240
Qy 241 CGTGISTRVNDNPECLVKETRICVVRPCGQPVYSSLKKGKCKSKTKKSPVPVFTYAG 300
Db 241 CGTGISTRVNDNPECLVKETRICVVRPCGQPVYSSLKKGKCKSKTKKSPVPVFTYAG 300
Qy 301 CSSVKYRYPKYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFKNVMMIQSKCNYNCPHP 360
Db 301 CSSVKYRYPKYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFKNVMMIQSKCNYNCPHP 360
Qy 361 NEASFRLYSLFNDIHKPRD 379
Db 361 NEASFRLYSLFNDIHKPRD 379

RESULT 2
US-09-142-569-2
; Sequence 2, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "Mouse Cyr61 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-142-569-2

Query Match 100.0%; Score 2103; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 5e-170;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPCLEAPKCAPGVGLVRDGGCGCKVCAKQL 60
Db 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPCLEAPKCAPGVGLVRDGGCGCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCCEVNSRIYQNGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCCEVNSRIYQNGESFQPNCKHQ 120
Qy 121 CTICIDGAVGCIPLCPQELSPLNLCNPNRPLVKVSGGCCBEEWDCDESIKSLDDDDLLG 180
Db 121 CTICIDGAVGCIPLCPQELSPLNLCNPNRPLVKVSGGCCBEEWDCDESIKSLDDDDLLG 180
Qy 181 LDASEVELTNNELIATGKSSLKRLPVGTEPRVLFNPLHAHQKCIQVTTTWSQCCKS 240
Db 181 LDASEVELTNNELIATGKSSLKRLPVGTEPRVLFNPLHAHQKCIQVTTTWSQCCKS 240
Qy 241 CGTGISTRVNDNPECLVKETRICVVRPCGQPVYSSLKKGKCKSKTKKSPVPVFTYAG 300
Db 241 CGTGISTRVNDNPECLVKETRICVVRPCGQPVYSSLKKGKCKSKTKKSPVPVFTYAG 300
Qy 301 CSSVKYRYPKYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFKNVMMIQSKCNYNCPHP 360
Db 301 CSSVKYRYPKYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFKNVMMIQSKCNYNCPHP 360
Qy 361 NEASFRLYSLFNDIHKPRD 379
Db 361 NEASFRLYSLFNDIHKPRD 379

RESULT 3
US-09-348-815-2
; Sequence 2, Application US/09348815
; Patent No. 6534630
; GENERAL INFORMATION:
; APPLICANT: LI, HAODONG
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/348,815
; FILING DATE: 08-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: JONATHAN L. KLEIN
; REGISTRATION NUMBER: 41,119
; REFERENCE/DOCKET NUMBER: PF126P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-348-815-2

Query Match 92.2%; Score 1938; DB 4; Length 381;

Best Local Similarity 91.4%; Pred. No. 4.5e-156;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFRFLAVAVTLLHLTRLALSTCPAACHPCLEAPKCAPGVGLVRDGGCCCKVC AKQL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHPCLEAPKCAPGVGLVRDGGCCCKVC AKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPOELSLNLCNPNRLVKVSGQCCCEWVCDDESIKSLDDDDLL-- 178
DB 121 CTCIDGAVGCIPLCPOELSLNLCNPNRLVKVSGQCCCEWVCDDESIKSLDDDDLL-- 178

QY 179 --LGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLFNPLHAGQKCIIVQTTSWSQ 236
DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLFNPLHAGQKCIIVQTTSWSQ 238

QY 237 CSKSGTGTSTRVNDNPECLRVKTRICEVRPCQGPVYSSLKGGKCKSKTKKSPPEVPVF 296
DB 239 CSKSGTGTSTRVNDNPECLRVKTRICEVRPCQGPVYSSLKGGKCKSKTKKSPPEVPVF 298

QY 297 TYAGCSSVKYRKYPCGSCVDGRCTPLOTTRVXMRFRCEDEGMFSKNVMMIQSKCKNYN 356
DB 299 TYAGCSSVKYRKYPCGSCVDGRCTPLOTTRVXMRFRCEDEGMFSKNVMMIQSKCKNYN 358

QY 357 CPHNEASFRLYSLFNDIHKFRD 379
DB 359 CPHANEAAFPFVRLFNDIHKFRD 381

RESULT 4

US-09-142-569-4
; Sequence 4, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Iau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: "Human Cyr61 amino acid sequence"

; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-142-569-4

Query Match 91.7%; Score 1929; DB 4; Length 381;
Best Local Similarity 90.9%; Pred. No. 2.6e-155;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFRFLAVAVTLLHLTRLALSTCPAACHPCLEAPKCAPGVGLVRDGGCCCKVC AKQL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHPCLEAPKCAPGVGLVRDGGCCCKVC AKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCYNRIYONGESFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPOELSLNLCNPNRLVKVSGQCCCEWVCDDESIKSLDDDDLL-- 178
DB 121 CTCIDGAVGCIPLCPOELSLNLCNPNRLVKVSGQCCCEWVCDDESIKSLDDDDLL-- 178

QY 179 --LGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLFNPLHAGQKCIIVQTTSWSQ 236
DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLFNPLHAGQKCIIVQTTSWSQ 238

QY 237 CSKSGTGTSTRVNDNPECLRVKTRICEVRPCQGPVYSSLKGGKCKSKTKKSPPEVPVF 296
DB 239 CSKSGTGTSTRVNDNPECLRVKTRICEVRPCQGPVYSSLKGGKCKSKTKKSPPEVPVF 298

QY 297 TYAGCSSVKYRKYPCGSCVDGRCTPLOTTRVXMRFRCEDEGMFSKNVMMIQSKCKNYN 356
DB 299 TYAGCSSVKYRKYPCGSCVDGRCTPLOTTRVXMRFRCEDEGMFSKNVMMIQSKCKNYN 358

QY 357 CPHNEASFRLYSLFNDIHKFRD 379
DB 359 CPHANEAAFPFVRLFNDIHKFRD 381

RESULT 5

US-08-468-847B-12
; Sequence 12, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:

LENGTH: 374 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-12

Query Match 86.5%; Score 1819.5; DB 1; Length 374;
Best Local Similarity 88.3%; Pred. No. 4.7e-146;
Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTRTLAVAVTLLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIVRELAVVTLHLTRVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFPQNCXHQ 120
DB 61 NEDCRKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFPQNCXHQ 120
QY 121 CTCIDGAVG-CIPLCPQELSLNLCNPGNPRLVKVSQCCEWVCDSDSIKDDDDDL- 178
DB 121 CTCIGWRRGACIPLCPQELSLNLCNPGNPRLVKVSQCCEWVCDSDSIKDDDDDL 180
QY 179 ---LGLDASEVELTRNNELIATGKSSLKELPVFGTEPRVLFNPLHAHQCKIVQTTWS 235
DB 181 GKLGLDASEVELTRNNELIATGKSSLKELPVFGTEPRVLFNPLHAHQCKIVQTTWS 238
QY 236 QCSKSGTGTSTRVNDNPECLVETRIKTRICEVPCGQPVYSLKKGKCKSKTKKSPPEVR 295
DB 239 QCSKSGTGTSTRVNDNPECLVETRIKTRICEVPCGQPVYSLKKGKCKSKTKKSPPEVR 298
QY 296 FTYAGSSVKYRKYPCGSCVDGRCTPLQTRTVKVRFCEDGEMFSKNVMIQSKCNY 355
DB 299 FTYAGCLSVKYRKYPCGSCVDGRCTPLQTRTVKVRFPCEDEGTFSKNVMIQSSKCN 358
QY 356 NCPHNEASFRLYSLF 371
DB 359 NCPHNEAAPPFYRLF 374

RESULT 6

US-08-459-101A-2
Sequence 2, Application US/08459101A
Patent No. 5945300
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Connective Tissue Growth Factor-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,101A
FILING DATE: June 2, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07736
FILING DATE: 12 JUL 94
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-317
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-459-101A-2

Query Match 86.2%; Score 1812.5; DB 2; Length 375;
Best Local Similarity 88.0%; Pred. No. 1.8e-145;
Matches 331; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

QY 1 MSSSTRTLAVAVTLLHLTRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSRIVRELAVVTLHLTRVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFPQNCXHQ 120
DB 61 NEDCRKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFPQNCXHQ 120
QY 121 CTCIDGAVG-CIPLCPQELSLNLCNPGNPRLVKVSQCCEWVCDSDSIKDDDDDL- 178
DB 121 CTCIGWRRGACIPLCPQELSLNLCNPGNPRLVKVSQCCEWVCDSDSIKDDDDDL 180
QY 179 ---LGLDASEVELTRNNELIATGKSSLKELPVFGTEPRVLFNPLHAHQCKIVQTTWS 235
DB 181 GKLGLDASEVELTRNNELIATGKSSLKELPVFGTEPRVLFNPLHAHQCKIVQTTWS 238
QY 236 QCSKSGTGTSTRVNDNPECLVETRIKTRICEVPCGQPVYSLKKGKCKSKTKKSPPEVR 295
DB 239 QCSKSGTGTSTRVNDNPECLVETRIKTRICEVPCGQPVYSLKKGKCKSKTKKSPPEVR 298
QY 296 FTYAGSSVKYRKYPCGSCVDGRCTPLQTRTVKVRFCEDGEMFSKNVMIQSKCNY 355
DB 299 FTYAGCLSVKYRKYPCGSCVDGRCTPLQTRTVKVRFPCEDEGTFSKNVMIQSSKCN 358
QY 356 NCPHNEASFRLYSLF 371
DB 359 NCPHNEAAPPFYRLF 374

RESULT 7

US-08-468-847B-13
Sequence 13, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

```

; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-13

Query Match      78.3%; Score 1646.5; DB 1; Length 375;
Best Local Similarity 80.1%; Pred. No. 2e-131;
Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;

QY 10 AVAVTLHLTRIAL-STCPAAACHPCLEAPKCAPGVGLVDRDGGCKKCAKQLNEDCSKTQ 68
Db 9 ALAAALLCLARLALGSPCAVQCQCPAAAPQCAPGVGLVDRDGGCKKCAKQLNEDCSKTQ 68
QY 69 PCDHKTGLECNFGASSTALKGICRAQSEGRPCVEYNSRIYQNGESFQPNCKHOCTCIDGAV 128
Db 69 PCDHKTGLECNFGASSTALKGICRAQSEGRPCVEYNSRIYQNGESFQPNCKHOCTCIDGAV 128
QY 129 GCIPLCPOELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL-----LGLDAS 184
Db 129 GCIPLCPOELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL-----LGLDAS 184
QY 185 EVELTRNNELIAGKSLKRLPVGTGP--RVLEPNFHAHQKCIQVTTTSWCSKSG 242
Db 187 EVELTRNNELIAGKSLKRLPVGTGP--RVLEPNFHAHQKCIQVTTTSWCSKSG 239
QY 243 TGISTRVTNDNPECLVETRIKCEVRPCGQPVYSSLLKGGKCKTKKSPSPVRYTACGS 302
Db 240 TGISTRVTNDNPECLVETRIKCEVRPCGQPVYSSLLKGGKCKTKKSPSPVRYTACGS 299
QY 303 SVKRYRPKYCGSVDGRCTPLQTRTVKMFRCEDGEMFSKNVMMIOQCKNVCNCPHNE 362
Db 300 SVKRYRPKYCGSVDGRCTPLQTRTVKMFRCEDGEMFSKNVMMIOQCKNVCNCPHNE 359
QY 363 ASFRLYSLFNDIHKPRD 379
Db 360 A-IPFRLVNDIHKPRD 375

RESULT 8
US-09-582-337-2
; Sequence 2, Application US/09582337
; Patent No. 6582619
; GENERAL INFORMATION:
; APPLICANT: Japan Tobacco, Inc.
; TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor
; FILE REFERENCE: J1-009PCT
; CURRENT APPLICATION NUMBER: US/09/582,337
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: JP 1997-367699
; PRIOR FILING DATE: 1997-12-25
; PRIOR APPLICATION NUMBER: JP 1998-356183
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Rat
US-09-582-337-2

Query Match      45.6%; Score 958; DB 4; Length 347;
Best Local Similarity 46.4%; Pred. No. 3.1e-73;
Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTRTTLLAVATLHL-TRIAL-STCPAAACHPCLE-APKCAPGVGLVDRDGGCKKCAK 57
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCCAAAPHPAGVSLVLDGGCCRCVCA 60
QY 58 KQLNEDCSKTQPCDHKTGLECNFGASSTALKGICRAQSEGRPCVEYNSRIYQNGESFQPNCK 117
Db 61 KQLELCTERDPCDPKGLFCDFGSPANRKGIVCTAK-DGAPCVFGSVYRSGESFQSSC 119
QY 118 KQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 177
Db 120 KYQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 168

```

```

Matches 176; Conservative 61; Mismatches 104; Indels 38; Gaps 7;

QY 1 MSSSTRTTLLAVATLHL-TRIAL-STCPAAACHPCLE-APKCAPGVGLVDRDGGCKKCAK 58
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCCAAAPHPAGVSLVLDGGCCRCVCA 60
QY 59 QLNEDCSKTQPCDHKTGLECNFGASSTALKGICRAQSEGRPCVEYNSRIYQNGESFQPNCK 118
Db 61 QLGELCTERDPCDPKGLFCDFGSPANRKGIVCTAK-DGAPCVFGSVYRSGESFQSSC 119
QY 119 HQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 178
Db 120 YQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 167
QY 179 LGLDASEVELTRNNELIAGKSLKRL-PVFGTGPVRLFNELFHAHQKCIQVTTTSWCS 237
Db 168 -----KORTVVGPALAAVRLDTFGDPDTM-----RANCLVQTTTSWCS 207
QY 238 SKSCGTGISTRVTNDNPECLVETRIKCEVRPCGQPVYSSLLKGGKCKTKKSPSPVRYT 297
Db 208 SKTCGMGISTRVTNDNPECLVETRIKCEVRPCGQPVYSSLLKGGKCKTKKSPSPVRYT 267
QY 298 YAGCSVKYRPKYCGSVDGRCTPLQTRTVKMFRCEDGEMFSKNVMMIOQCKNVCN 357
Db 268 LSGCTSVKTYRAKFCGCTDGRCTPHRTTLVPEKCPDGEIMKQNMFEIKTCAHYNC 327
QY 358 PHPNEASFRLY--SLFNDI 374
Db 328 PGDNDFESLYRYKMYGDM 346

RESULT 9
US-09-292-036-3
; Sequence 3, Application US/09292036
; Patent No. 6358741
; GENERAL INFORMATION:
; APPLICANT: FIBROGEN, INC
; APPLICANT: SCHMIDT, Brian
; APPLICANT: ALLEN, Margaret
; APPLICANT: SVERDRUP, Fran
; APPLICANT: CARMICHAEL, David
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE
; FILE REFERENCE: FIBRO1100-1
; CURRENT APPLICATION NUMBER: US/09/292,036
; CURRENT FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/292,036
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/187,478
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Mouse
US-09-292-036-3

Query Match      45.5%; Score 957.5; DB 4; Length 348;
Best Local Similarity 46.8%; Pred. No. 3.4e-73;
Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTRTTLLAVATLHL-TRIAL-STCPAAACHPCLE-APKCAPGVGLVDRDGGCKKCAK 57
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCCAAAPHPAGVSLVLDGGCCRCVCA 60
QY 58 KQLNEDCSKTQPCDHKTGLECNFGASSTALKGICRAQSEGRPCVEYNSRIYQNGESFQPNCK 117
Db 61 KQLELCTERDPCDPKGLFCDFGSPANRKGIVCTAK-DGAPCVFGSVYRSGESFQSSC 119
QY 118 KQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 177
Db 120 KYQCTCIDGAVGICPILCPQELSLPNLGCNPNRLVKVSGQCCBEWVDEDSIKDSDLDQDDL 168

```

QY 178 LLGLDASEVELTRNNELIALGKSSIKRL-PVFGTEPRVLFNPLHAHGQKCIQVTTSSQ 236
Db 169 -----KORTAVGPALAAAYRLDTFGDPDTM-----RANCLVQTTWSA 207
QY 237 CSKSCGTGISTRTVNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 296
Db 208 CSKTCGMGISTRTVNDNTEFCRLKQSLCWLVRPCEADLEENIKKGGKCIKRTPKIAKPVKF 267
QY 297 TYAGCSVKYRKYCGSCVDGRCCCTPLOTTRIVKMRFRCEDEMGEMSKNVMIOCKKCNYN 356
Db 268 ELSGCTSVKTYRAKFCGVCCTDGRCCCTPHRTTTLVPEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPGDNDIFESLYRKMVGDM 347

RESULT 10

US-08-468-847B-15
; Sequence 15, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human Ccn-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-15

Query Match 45.3%; Score 953.5; DB 1; Length 348;

Best Local Similarity 46.6%; Pred. No. 7.5e-73;

Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFRTLAVAVTLHL-TRLAL-STCPAACHPCLE-APKCAPGVGVLDGCGCKVCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGDCSAQCCAAEAAPHCPAGVSLVLDGCGCCVCA 60
QY 58 KQLNEDCSKTPQCDHTKGLCNFGASSTALKGCRAQSEGRFCPEYNSRIYQNGESFPQNC 117
Db 61 KQLGELCTERDPCDPKHLFCDFGSPANRKGIVCTAK-DGAPCVFGSVYRSGESFQSSC 119

QY 118 KHQCTCIDGAVGCIPLCPQBELSLPNLCNPNRLVKVSGQCCWEVWVDEDSIKDSDDD 177
Db 120 KYQCTCLDGVAVGVPLCSMDVRLPSPCPPEPRVRLKPGCKCKEWVCDPE 168
QY 178 LLGLDASEVELTRNNELIALGKSSIKRL-PVFGTEPRVLFNPLHAHGQKCIQVTTSSQ 236
Db 169 -----KORTAVGPALAAAYRLDTFGDPDTM-----RANCLVQTTWSA 207
QY 237 CSKSCGTGISTRTVNDNPECLVKETRICVVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 296
Db 208 CSKTCGMGISTRTVNDNTEFCRLKQSLCWLVRPCEADLEENIKKGGKCIKRTPKIAKPVKF 267
QY 297 TYAGCSVKYRKYCGSCVDGRCCCTPLOTTRIVKMRFRCEDEMGEMSKNVMIOCKKCNYN 356
Db 268 ELSGCTSVKTYRAKFCGVCCTDGRCCCTPHRTTTLVPEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPGDNDIFESLYRKMVGDM 347

RESULT 11

US-09-142-569-6
; Sequence 6, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: "Fisp12 amino acid sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-142-569-6

Query Match 45.3%; Score 953.5; DB 4; Length 348;

Best Local Similarity 46.6%; Pred. No. 7.5e-73;

Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFRTLAVAVTLHL-TRLAL-STCPAACHPCLE-APKCAPGVGVLDGCGCKVCA 57

Db 1 MLASVAGPISLALVLLALCTRPATGDCSAQCCAAEAAPHCPAGVSLVLDGCGCCVCA 60

QY 58 KOLNEDCKTOPCDHTKGLNCFNFGASSTALGICRAQSEGRPCPEYNSRIYONGESPOPNC 117
 DB 61 KOLGELCTERPCDPHKLFLCDFGSPANRKTIGVCTAK-DGAPCVFGSVYRSGESFOSSC 119
 QY 118 KHQCTCIDGAVGCIPLCPQELSLNLCGNPRLVKVSGQCEEWVCDSDSKDLDQDD 177
 DB 120 KYQCTCLDGVGCVPLCSMDVRLPSPDCPPFRRVKLPKCKCKEWVCDSP 168
 QY 178 LLGLDASEVELTRNNELIAIGKSSSLKRL-PVFGTEPRVLFNPLHAGQKCIYQTTWSQ 236
 DB 169 -----KRTAVGPAALAAVRLDTFGPDPTM-----RANCLVQTTWSA 207
 QY 237 CSKSGTGTSTRVNDNPECLVKETRICEVRPCGQPVYSSLKKGKCKSKTKSPVRF 236
 DB 208 CSKTCGMISTRVNDNTFCLEKQSLCMLVRPCEADLEENIKGKCKCIPTPKIAKVKF 267
 QY 297 TYAGSSVKYRKYCGSVDCRCCTPLOTQRTVVMRFRCEBDEGEMFSKNVMMIOSCKNYN 356
 DB 268 ELSGCTSVKTYRAKFCGVCCTDGRCTPHRTTTLPEVKCPDGEIMKNNMFITCACHYN 327
 QY 357 CPHNEASFRLY--SLENDI 374
 DB 328 CPGNDIFESLYYRMYGDM 347

RESULT 12
 US-08-167-628-2
 ; Sequence 2, Application US/08167628
 ; Patent No. 5408040
 ; GENERAL INFORMATION:
 ; APPLICANT: Grotendorst, Gary R.
 ; APPLICANT: Bradham Jr., Douglas M.,
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,628
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/752,427
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wetherell, Jr. Ph.D., John W.
 ; REGISTRATION NUMBER: 31,678
 ; REFERENCE/DOCKET NUMBER: PD-1294
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-455-5100
 ; TELEFAX: 619-455-5110
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 349 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-167-628-2

Query Match 45.2%; Score 950; DB 1; Length 349;
 Best Local Similarity 45.7%; Pred. No. 1.5e-72;
 Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRTLAVATLHLRLALSTCPAA-----CHCPLE-APKCAPGVGLVRDGGCC 52
 DB 1 MTAASMGPRVAFVVL-----LALCSRPAVGQNCSPGCPDEPAPRCRAGVSLVLDGGCC 56
 QY 53 CKVCAKQLNEDCKTOPCDHTKGLNCFNFGASSTALGICRAQSEGRPCPEYNSRIYONGES 112
 DB 57 CRVCAKQLGELCTERPCDPHKLFLCDFGSPANRKTIGVCTAK-DGAPCIFGTVYRSGES 115
 QY 113 FQPNCKHGOCTCIDGAVGCIPLCPQELSLNLCGNPRLVKVSGQCEEWVCDSDSKDLS 172
 DB 116 FOSSKRYQCTCLDGVGCVPLCSMDVRLPSPDCPPFRRVKLPKCKCKEWVCDSPK 170
 QY 173 DQDDLLGLDASEVELTRNNELIAIGKSSSLKRL-PVFGTEPRVLFNPLHAGQKCIYQTT 231
 DB 171 -DQ-----TVVGPALAAVRLDTFGPDPTM-----RANCLVQTT 203
 QY 232 TWSQCSKSGTGTSTRVNDNPECLVKETRICEVRPCGQPVYSSLKKGKCKSKTKKSP 291
 DB 204 TWSACSKTCGMISTRVNDNTFCLEKQSLCMLVRPCEADLEENIKGKCKCIPTPKIS 263
 QY 292 EPRVITYAGSSVKYRKYCGSVDCRCCTPLOTQRTVVMRFRCEBDEGEMFSKNVMMIOSC 351
 DB 264 KPIKELSGCTSVKTYRAKFCGVCCTDGRCTPHRTTTLPEVKCPDGEIMKNNMFITC 323
 QY 352 KNYNCPHNEASFRLY--SLENDI 374
 DB 324 ACHYNCPGNDIFESLYYRMYGDM 348

RESULT 13
 US-08-386-680-2
 ; Sequence 2, Application US/08386680
 ; Patent No. 5585270
 ; GENERAL INFORMATION:
 ; APPLICANT: Grotendorst, Gary R.
 ; APPLICANT: Bradham Jr., Douglas M.,
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/386,680
 ; FILING DATE: 10-FEB-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,628
 ; FILING DATE:
 ; APPLICATION NUMBER: US/07/752,427
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wetherell, Jr. Ph.D., John W.
 ; REGISTRATION NUMBER: 31,678
 ; REFERENCE/DOCKET NUMBER: PD-1294
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-455-5100
 ; TELEFAX: 619-455-5110
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 349 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-386-680-2

```

Query Match      45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.5e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRTAVATLHLTRLALSTCPAA-----CHCPLE-APKCAPGVGLVRDGGCG 52
Db 1 MTAASGPRVAVFVVL-----LALCSRPAVGQNCSPCRCPDEPAPRCFAGVSLVDGGCG 56
QY 53 CRVCAKQLNEDSKTOPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRIYQNGES 112
Db 57 CRVCAKQLGELCTERPCDPKGLFCDFGSPANRKGIVCTAK-DGAPCIFGGTVYRSGES 115
QY 113 FQNCXKQCTCIDGAGCTPLCPQELSLPNLGCNPNRLVKVSGCCCEWVDEDSIKDSL 172
Db 116 FOSSCKYQCTCLDAGVGMPLCSMDVRLPSPDCPPFRVRKLPKGCCEWVDEPK----- 170
QY 173 DDQDILLGLDASEVELTRNNELIAIGKSSSLKRL-PVFGTEPRVLFNPLHAHQKCIQVOT 231
Db 171 -DQ-----TVVGPALAAVRLDTFGDPDTMI-----RANCLVQT 203
QY 232 TWSQSKSGCGTIGSTRVNDNPECLVKEIRICEVRPCQGVYSSLKKGKCKTKKSP 291
Db 204 TWSACKSKTGMGISTRTVNDNASCRLKQSLCMVRPCEADLEENIKKGGKCIPTPKIS 263
QY 292 EPVRFYAGSSVKYKRYKPCGSDVGRCTPLQTRTVKMRPRCEDEGEMFSKNVMMIOSC 351
Db 264 KPIKFLSGCTSMKTYRAKFCGVCTDGRCTPHRTITLTVKFCPCDGEVMMKNMFIKTC 323
QY 352 KQNYNCPHNEASFRLY--SLFNDI 374
Db 324 ACHYNCFGDNDIFESLYRYKMYGDM 348

```

```

RESULT 14
US-08-459-717-2
; Sequence 2, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-717-2
Query Match      45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.5e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRTAVATLHLTRLALSTCPAA-----CHCPLE-APKCAPGVGLVRDGGCG 52
Db 1 MTAASGPRVAVFVVL-----LALCSRPAVGQNCSPCRCPDEPAPRCFAGVSLVDGGCG 56
QY 53 CRVCAKQLNEDSKTOPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRIYQNGES 112
Db 57 CRVCAKQLGELCTERPCDPKGLFCDFGSPANRKGIVCTAK-DGAPCIFGGTVYRSGES 115
QY 113 FQNCXKQCTCIDGAGCTPLCPQELSLPNLGCNPNRLVKVSGCCCEWVDEDSIKDSL 172
Db 116 FOSSCKYQCTCLDAGVGMPLCSMDVRLPSPDCPPFRVRKLPKGCCEWVDEPK----- 170
QY 173 DDQDILLGLDASEVELTRNNELIAIGKSSSLKRL-PVFGTEPRVLFNPLHAHQKCIQVOT 231
Db 171 -DQ-----TVVGPALAAVRLDTFGDPDTMI-----RANCLVQT 203
QY 232 TWSQSKSGCGTIGSTRVNDNPECLVKEIRICEVRPCQGVYSSLKKGKCKTKKSP 291
Db 204 TWSACKSKTGMGISTRTVNDNASCRLKQSLCMVRPCEADLEENIKKGGKCIPTPKIS 263
QY 292 EPVRFYAGSSVKYKRYKPCGSDVGRCTPLQTRTVKMRPRCEDEGEMFSKNVMMIOSC 351
Db 264 KPIKFLSGCTSMKTYRAKFCGVCTDGRCTPHRTITLTVKFCPCDGEVMMKNMFIKTC 323
QY 352 KQNYNCPHNEASFRLY--SLFNDI 374
Db 324 ACHYNCFGDNDIFESLYRYKMYGDM 348

RESULT 15
US-08-712-302-2
; Sequence 2, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,302
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:51:09 ; Search time 49.8684 Seconds
(without alignments)
2147.361 Million cell updates/sec

Title: US-09-495-448A-2
Perfect score: 2103
Sequence: 1 MSSSTFRFLAVVILLHTR.....PNEASFRLYSLFNDHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	2	AAR25565 Beta-IG-M
2	2103	100.0	379	4	AAR25565 Mouse cys
3	2103	100.0	379	5	ABO09201 Mouse cys
4	1938	92.2	381	2	AAR35957 Human con
5	1938	92.2	381	4	AAB390773 Human she
6	1938	92.2	381	5	AAR79761 Human cys
7	1938	92.2	381	5	ABO05438 Human cys
8	1938	92.2	381	5	ABO05438 Human cys
9	1938	92.2	381	6	ABU63222 Human con
10	1938	92.2	455	3	AAB43987 Human can
11	1930	91.8	381	5	ABG76937 Human pro
12	1930	91.8	381	7	ADB75260 Prostate
13	1929	91.7	381	2	AAR35730 Human cys
14	1929	91.7	381	4	AAR05921 Human cys
15	1819.5	86.5	374	5	ABO09202 HCGF CNN
16	1819.5	86.5	375	2	AAR90919 Connectiv
17	1819.5	86.5	375	2	AAY31620 Human CTG
18	1819.5	86.5	375	5	ABU63223 Human alt
19	1819.5	86.5	375	6	ABU63223 Human con
20	1646.5	78.3	375	4	AAR05939 Chicken c
21	1646.5	78.3	375	5	ABO09203 Chicken c
22	960	45.6	347	6	ABR43138 Rat conne
23	958	45.6	347	2	AAY24379 Rat conne
24	957.5	45.5	348	2	AAR25566 Beta-IG-M
25	957.5	45.5	348	6	ADB25766 Mouse con

ALIGNMENTS

RESULT 1
AAR25565
ID AAR25565 standard; protein; 379 AA.

AC AAR25565;
XX
XX 25-MAR-2003 (revised)
DT 18-JAN-1993 (first entry)
XX
XX
DB Beta-IG-M1.

XX Transforming growth factor beta; induced; CEF-10; v-src; chicken; embryo;
KW Fibroblasts; TGF-beta.
XX
XX Mus musculus.

OS
XX
XX EP495674-A2.
PN
XX
XX 22-JUL-1992.

XX 17-JAN-1992; 92EP-00300429.
PF
XX
XX 18-JAN-1991; 91US-00642991.
PR 10-JAN-1992; 92US-00816270.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Purchio AP, Brunner AM, Chinn J, Neubauer MG;
PI
XX WPI; 1992-243508/30.
DR
XX N-PSDB; AAQ26421.

XX TGF-beta induced gene family - encodes proteins involved in growth and
PT differentiation effects of TGF-beta-1.
XX
XX Claim 2; Fig 1; 35pp; English.

CC The protein sequence was deduced from the DNA sequence obtd. by screening
CC a cDNA library made from AKR-2B mouse cells induced with TGF-beta1 and
CC cyclohexamide with two probes from untreated AKR-2B mRNA and AKR-2B mRNA
CC from cells treated with cyclohexamide and TGF-beta1. The proteins
CC encoded by hybridising colonies (beta-IG-M1 and beta-IG-M2) contain 38
CC Cys residues and are induced by TGF-beta1. Beta-IG-M1 displays 80 percent
CC homology to the CEF-10 protein induced by v-src in chicken embryo
CC fibroblasts and is identical to the protein encoded by cyr51, an
CC immediate early response gene induced in quiescent BALB 3T3 cells by
CC serum treatment. Residues 49-56 of beta-IG-M1 conform to the GCGCCXC
CC motif reported in the amino half of insulin-like growth factor (IGF)

ADB25762 Mouse con
Aaw35731 Murine fi
Aay44756 Mouse con
Aae05922 Mouse fib
Aab09205 Fisp-12 C
Abr43139 Mouse con
Aaw12694 Connectiv
Aar79964 Connectiv
Aaw11302 Connectiv
Aaw09089 Human con
Aaw2084 Human con
Aaw81425 Connectiv
Aay18361 Human con
Aay92939 Human con
Aay92940 Human con
Aay44755 Amino aci
Aab44598 Amino aci
Aab07991 Human she
Aab06664 Human con
Aab48831 Human con

CC binding proteins. The C-terminal Cys rich region of beta-IG-M1, -M2 and
 CC CSE-10 contain an amino acid sequence with strong homology to a motif
 CC found near the C-terminal of the malarial circumsporozoite (CS) protein,
 CC which is highly conserved among all species of malarial parasites
 CC sequenced to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate cell-cell
 CC and cell-extracellular matrix interactions, such as properdin,
 CC thrombospondin, and TRAP. The proteins encoded by TGF-beta induced genes
 CC are likely to be involved in mediation of the biological effects of TGF-
 CC beta relating to cell growth and differentiation. See also AAR25566.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 379 AA;

Query Match 100.0%; Score 2103; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 2.7e-155;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSSTFRTLAVAVTLLHLTLALSTCPAACHCPLKAPGVGLVDRDGGCKVCVKAKQL 60
 DB 1 MSSSTFRTLAVAVTLLHLTLALSTCPAACHCPLKAPGVGLVDRDGGCKVCVKAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCPEYNSRIYQNGESFPQNCCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCPEYNSRIYQNGESFPQNCCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGQCEEWVCDSDSIKSLDDQDLDLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGQCEEWVCDSDSIKSLDDQDLDLLG 180
 QY 181 LDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLNPLHAGQKCIQVTTSSWSQCSKS 240
 DB 181 LDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLNPLHAGQKCIQVTTSSWSQCSKS 240
 QY 241 CGTGISTRTVNDNPECLVKETRICVVRPGQPVYSSLKKGKCKTKKSPPEVRFTYAG 300
 DB 241 CGTGISTRTVNDNPECLVKETRICVVRPGQPVYSSLKKGKCKTKKSPPEVRFTYAG 300
 QY 301 CSSVKKYRKYKCGSDVGRCTPLQTRTVKMRFRCEDEGMFSSKNVMMIQSKCNYNCPHP 360
 DB 301 CSSVKKYRKYKCGSDVGRCTPLQTRTVKMRFRCEDEGMFSSKNVMMIQSKCNYNCPHP 360
 QY 361 NEASFRLYSFLNDIHKFRD 379
 DB 361 NEASFRLYSFLNDIHKFRD 379

RESULT 2
 AAE05920
 ID AAE05920 standard; protein; 379 AA.
 XX
 AC AAE05920;
 XX
 DT 24-SEP-2001 (first entry)
 XX

DE Mouse cysteine-rich protein (Cyr61).
 XX
 KW Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KW fibroblast secreted protein; Fisp12; connective tissue growth factor;
 KW CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 KW angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KW heart disease; fibrosis; gene therapy; mouse.
 XX

OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Region 164..226
 FT /note= "Cysteine free region"
 FT Domain 224..240
 FT /note= "Domain III"
 XX
 FN WO200155210-A2.
 XX

PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US003267.
 XX
 PR 31-JAN-2000; 2000US-00495448.
 PR 15-MAY-2000; 2000US-0204364P.
 PR 06-OCT-2000; 2000US-0238705P.
 XX
 PA (MUNI-) MUNIN CORP.
 XX
 PI Lau LF, Yeung C, Greenspan JA;
 XX
 XX WPI; 2001-465561/50.
 DR N-PSDB; AAD11220.
 XX
 PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration.
 XX
 XX Example 1; Fig 1; 186pp; English.

CC The invention relates to extracellular matrix (ECM) signalling molecules
 CC involved in cellular response to growth factors. More particularly the
 CC invention is directed to cysteine-rich protein (Cyr61), and Cyr61-related
 CC proteins such as fibroblast secreted protein (Fisp12), and connective
 CC tissue growth factor (CTGF) and nucleic acid molecules encoding such
 CC proteins. The polypeptides of the invention are members of cysteine-rich
 CC secreted protein family. Human Cyr61 fragment is useful in methods for
 CC screening modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin receptor
 CC interaction. Modulator of Cyr61-integrin alphavbeta3 interaction is used
 CC for the preparation of a medicament for the treatment of atherosclerosis,
 CC heart disease, tumour metastasis, fibrosis, tumour growth, disorders
 CC associated with inadequate angiogenesis; aberrant granulation tissue
 CC development; aberrant fibroblast growth and wounds. Polynucleotides of
 CC the invention are useful in gene therapy. The present sequence is mouse
 CC Cyr61 protein
 XX

SQ Sequence 379 AA;

Query Match 100.0%; Score 2103; DB 4; Length 379;
 Best Local Similarity 100.0%; Pred. No. 2.7e-155;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSSTFRTLAVAVTLLHLTLALSTCPAACHCPLKAPGVGLVDRDGGCKVCVKAKQL 60
 DB 1 MSSSTFRTLAVAVTLLHLTLALSTCPAACHCPLKAPGVGLVDRDGGCKVCVKAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCPEYNSRIYQNGESFPQNCCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASTALKGICRAQSEGRPCPEYNSRIYQNGESFPQNCCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGQCEEWVCDSDSIKSLDDQDLDLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGQCEEWVCDSDSIKSLDDQDLDLLG 180
 QY 181 LDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLNPLHAGQKCIQVTTSSWSQCSKS 240
 DB 181 LDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLNPLHAGQKCIQVTTSSWSQCSKS 240
 QY 241 CGTGISTRTVNDNPECLVKETRICVVRPGQPVYSSLKKGKCKTKKSPPEVRFTYAG 300
 DB 241 CGTGISTRTVNDNPECLVKETRICVVRPGQPVYSSLKKGKCKTKKSPPEVRFTYAG 300
 QY 301 CSSVKKYRKYKCGSDVGRCTPLQTRTVKMRFRCEDEGMFSSKNVMMIQSKCNYNCPHP 360
 DB 301 CSSVKKYRKYKCGSDVGRCTPLQTRTVKMRFRCEDEGMFSSKNVMMIQSKCNYNCPHP 360
 QY 361 NEASFRLYSFLNDIHKFRD 379
 DB 361 NEASFRLYSFLNDIHKFRD 379

RESULT 3
 ABB09201
 ID ABB09201 standard; protein; 379 AA.
 XX
 AC ABB09201;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Mouse cyr6 CNN family protein sequence SEQ ID NO:11.
 XX
 KW Human; small CCN-like growth factor; SCGF; vulnery; osteopathic;
 KW Gene therapy; muscle wasting disease; osteoporosis; wound healing;
 KW tissue regeneration; angiogenesis.
 XX
 OS Mus sp.
 XX
 PN US2002049304-A1.
 XX
 PD 25-APR-2002.
 XX
 PF 14-MAY-2001; 2001US-00853625.
 XX
 PR 06-JUN-1995; 95US-00468847.
 PR 01-APR-1998; 98US-00053587.
 XX
 PA (HAST/) HASTINGS G A.
 PA (ADAM/) ADAMS M D.
 XX
 PI Hastings GA, Adams MD;
 XX
 DR WPI; 2002-382150/41.
 XX
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 PT growth factor, useful for treating muscle wasting disease, and
 PT osteoporosis.
 XX
 PS Disclosure; Fig 2A-D; 33pp; English.
 XX
 CC The present invention describes human small CCN-like growth factor
 CC (SCGF). SCGF has vulnery and osteopathic activities, and can be used in
 CC gene therapy. The SCGF polypeptides and polynucleotides can be used for
 CC treating muscle wasting diseases, and osteoporosis, and to stimulate
 CC wound healing and tissue regeneration, to promote angiogenesis and to
 CC stimulate proliferation of vascular smooth muscle and endothelial cell
 CC production. The present sequence represents a CNN family protein which is
 CC given in comparison with the human SCGF in the exemplification of the
 CC present invention
 XX
 SQ Sequence 379 AA;
 Query Match 100.0%; Score 2103; DB 5; Length 379;
 Best Local Similarity 100.0%; Pred. No. 2.7e-155;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
 DB 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYNGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCQELSLNLCGNPRLVKVSGCCCEWVCDSDIKSLDDDDLLG 180
 DB 121 CTCIDGAVGCIPLCQELSLNLCGNPRLVKVSGCCCEWVCDSDIKSLDDDDLLG 180
 QY 181 LDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLFNPLHAHQKCIQVOTTSWSQSKS 240
 DB 181 LDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLFNPLHAHQKCIQVOTTSWSQSKS 240
 QY 241 CGTGISTVTNDNPECRVLKTRICEVRPCQGPVYSSLLKKGCKSKTKKSPVPVRYAG 300
 DB 241 CGTGISTVTNDNPECRVLKTRICEVRPCQGPVYSSLLKKGCKSKTKKSPVPVRYAG 300

QY 301 CSSVKYRKYCGSCVDGRCCTPLQTRTVKMFRCEDGEMFSKNVMMIQSKCNYNCPHP 360
 DB 301 CSSVKYRKYCGSCVDGRCCTPLQTRTVKMFRCEDGEMFSKNVMMIQSKCNYNCPHP 360
 QY 361 NEASFLYSLFNDIHKFRD 379
 DB 361 NEASFLYSLFNDIHKFRD 379
 RESULT 4
 AAW35957
 ID AAW35957 standard; protein; 381 AA.
 XX
 AC AAW35957;
 XX
 DT 05-MAR-1998 (first entry)
 XX
 DE Human monocyte mature differentiation factor.
 XX
 KW Human; monocyte; mature; differentiation factor; MMDF; macrophage;
 KW cancer; immune activator; tissue culture; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN JP09234079-A.
 XX
 PD 09-SEP-1997.
 XX
 PF 04-MAR-1996; 96JP-00075236.
 XX
 PR 04-MAR-1996; 96JP-00075236.
 XX
 PA (TOYM) TOYBO KK.
 XX
 DR WPI; 1997-497320/46.
 DR N-PSDB; AAT97142.
 XX
 XX A monocyte mature differentiation factor - useful for the long term
 PT tissue culture of macrophage(s).
 XX
 PS Claim 9; Page 12-13; 22pp; Japanese.
 CC
 CC The present sequence represents a monocyte mature differentiation factor
 CC (MMDF) which maintains the life of macrophages for long periods in liquid
 CC culture. MMDF can be used as an anti-cancer agent, an immune activator
 CC and to treat infectious diseases
 XX
 SQ Sequence 381 AA;
 Query Match 92.2%; Score 1938; DB 2; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.9e-142;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSSTFRTLAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
 DB 1 MSSRTARALAVVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYNGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCQELSLNLCGNPRLVKVSGCCCEWVCDSDIKSLDDDDLLG 178
 DB 121 CTCIDGAVGCIPLCQELSLNLCGNPRLVKVSGCCCEWVCDSDIKSLDDDDLLG 180
 QY 179 --LGDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLFNPLHAHQKCIQVOTTSWSQ 236
 DB 181 KELGFDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLFNPL--QGCKCIVOTTSWSQ 238
 QY 237 CSKSCGTGISTVTNDNPECRVLKTRICEVRPCQGPVYSSLLKKGCKSKTKKSPVPVRY 296
 DB 239 CSKTCGTGISTVTNDNPECRVLKTRICEVRPCQGPVYSSLLKKGCKSKTKKSPVPVRY 298

QY 297 TYAGCSSVKYRKYKCGSCVDGRCCCTPLOTRTVYMRFRCEGEMFSKNVMMIOSCKCNYN 356
 Db 299 TYAGCLSVKRYRKYKCGSCVDGRCCCTPLOTRTVYMRFRCEGEMFSKNVMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 Db 359 CPHANEAAFPFYELFNDIHKFRD 381

RESULT 5
 AAB90773
 ID AAB90773 standard; protein; 381 AA.
 XX AAB90773;
 XX 15-JUN-2001 (first entry)
 DT Human shear stress-response protein SEQ ID NO: 46.
 DE Human; shear stress-response protein; vascular disease; arteriosclerosis.
 KW Human; shear stress-response protein; vascular disease; arteriosclerosis.
 OS Homo sapiens.
 XX WO200125427-A1.
 FN 12-APR-2001.
 XX 02-OCT-2000; 2000WO-JP006840.
 XX 01-OCT-1999; 99JP-00280976.
 PR (KYOW) KYOWA HAKKO KOGYO KK.
 XX (NOJII) NOJIMA H.
 PA Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K,
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 XX WPI; 2001-266308/27.
 DR N-PSDB; AAB02896.
 XX DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis.
 XX Claim 60; Page 345-346; 678pp; Japanese.
 PS The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension
 XX Sequence 381 AA;
 SQ

Query Match 92.2%; Score 1938; DB 4; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.9e-142;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTRTLAVATLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCGCKVCAKQL 60
 Db 1 MSSRIARALAVTLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCGCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCENSRVYQNGESFQNCXHQ 120
 Db 61 NEDCSKTQPCDHTKGLNCFNFGASSTALKGICRAQSEGRPCENSRVYQNGESFQNCXHQ 120
 QY 121 CTCIDGAVGICPLCPQELSLPGLNCGPNRLVKVSCCEEWVCDSDSIKDSLDODDL-- 178
 Db 121 CTCIDGAVGICPLCPQELSLPGLNCGPNRLVKVSCCEEWVCDSDSIKDPMEDQDGLG 180
 QY 179 --LGLDASEVELTRNELLIAIGKSSLRKLPVFGTEPRVLFNPLHAHGOKICIVOTTSWSQ 236

Db 181 KEIGFDASEVELTRNELLIAVGKSSLRKLPVFGTEPRVLFNPL--QGQKCIVOTTSWSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKETRICVREPCGQPVYSSLLKGGKCKSKTKSPBPVRF 236
 Db 239 CSKTCGTGISTRTVNDNPECLVKETRICVREPCGQPVYSSLLKGGKCKSKTKSPBPVRF 298
 QY 297 TYAGCSSVKYRKYKCGSCVDGRCCCTPLOTRTVYMRFRCEGEMFSKNVMMIOSCKCNYN 356
 Db 299 TYAGCLSVKRYRKYKCGSCVDGRCCCTPLOTRTVYMRFRCEGEMFSKNVMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 Db 359 CPHANEAAFPFYELFNDIHKFRD 381

RESULT 6
 AAU79761
 ID AAU79761 standard; protein; 381 AA.
 XX AAU79761;
 AC 30-JUL-2002 (first entry)
 DT Human Cyr61 protein.
 DE Human; uterine leiomyoma proliferation; uterine leiomyoma formation;
 XX Cyr61; cytotostatic.
 KW Homo sapiens.
 OS WO200226193-A2.
 FN 04-APR-2002.
 XX 28-SEP-2001; 2001WO-US030783.
 XX 29-SEP-2000; 2000US-0236887P.
 PR (AMHP) AMERICAN HOME PROD CORP.
 XX Zhang Z, Sampath D, Zhu Y, Minneker R;
 PI WPI; 2002-383245/41.
 DR N-PSDB; ABK48899.
 XX Preventing uterine leiomyoma formation or inhibiting proliferation of
 PT uterine leiomyoma in subject, comprises modulating or increasing the
 PT level of Cyr61 in leiomyoma tissue.
 XX Disclosure; Fig 6; 92pp; English.
 CC The present invention relates to a method of inhibiting proliferation of
 CC uterine leiomyoma or preventing uterine leiomyoma formation. The method
 CC comprises increasing the level of Cyr61 in leiomyoma tissue. The
 CC invention also describes compounds and compositions that stimulate
 CC induction of the Cyr61 gene and compounds that increase Cyr61 activity.
 CC The compositions and the method of the invention are useful for
 CC preventing uterine leiomyoma formation or inhibiting proliferation of
 CC uterine leiomyoma in a subject. The method is particularly useful for
 CC treating or preventing uterine leiomyoma formation, or inhibiting
 CC proliferation of uterine leiomyoma in a subject. The present sequence
 CC represents human Cyr61. Note: The present sequence shown in Fig 6 is not
 CC shown in the correct sequence order in the figure. The start of the
 CC sequence is shown on page 8/10 of the figures and the rest of the
 CC sequence is shown on page 6/10 of the figures
 XX Sequence 381 AA;
 SQ

Query Match 92.2%; Score 1938; DB 5; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.9e-142;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTRTLAVATLHLTRLALSTCPAACHCPLKAPGVGLVRDGGCGCKVCAKQL 60

Db 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCGCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENRIYONGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENRIYONGESFQPNCKHQ 120
Qy 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVSGQCEEWVCDSDIKSLDDDDLL-- 178
Db 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVSGQCEEWVCDSDIKSLDDDDLLG 180
Qy 179 --LGLDASEVELTRNNELIAGKSSLRKLPVFGTEPRVLPNPLHAHQKCIQVTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAGKSSLRKLPVFGTEPRVLPNPLHAHQKCIQVTTWSQ 238
Qy 237 CSKSGCTGISTRVTNDNPECLVKEVTRICEVRPCQPVYSSLRKGLKSKTKKSPPEVRF 296
Db 239 CSKSGCTGISTRVTNDNPECLVKEVTRICEVRPCQPVYSSLRKGLKSKTKKSPPEVRF 298
Qy 297 TYAGSSVKYRPRKYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNVMMIQSKCKNYN 356
Db 299 TYAGCLSVKYRPRKYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNVMMIQSKCKNYN 358
Qy 357 CPHPNEASFRLYSLFNDIHKFRD 379
Db 359 CPHPNEAAPPYRLFNDIHKFRD 381
RESULT 7
ABB05438
ID ABB05438 standard; protein; 381 AA.
XX ABB05438;
XX
DT 15-APR-2002. (first entry)
DE Human Cyr61 protein SEQ ID NO:2.
KW Human; Cyr61; breast cancer; sex steroid receptor; cytostatic; promoter;
KW sex steroid response element; cysteine rich heparin-binding protein;
KW cell proliferation; heparin binding epidermal growth factor;
KW epidermal growth factor; basic fibroblastic growth factor.
XX
OS Homo sapiens.
XX
XX WO200198359-A2.
XX
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-US019823.
XX
XX 21-JUN-2000; 2000US-0213182P.
XX
XX 16-MAY-2001; 2001US-0291510P.
XX
XX (AMHP) AMERICAN HOME PROD CORP.
XX
XX Sampath D, Zhang Z, Winneker R;
XX
XX WPI; 2002-147796/19.
XX
XX N-PSDB; ABA93127, ABA93130.
XX
XX Regulation of Cyr61 expression and activity for preventing and inhibiting
XX breast cancer comprises use of a Cyr61 neutralizing antibody, an anti-
XX sense oligonucleotide and an antibody which.
XX
XX Claim 6; Fig 1; 86pp; English.
XX
XX The present invention describes a method for the prevention or inhibition
XX of breast cancer cell proliferation. The method comprises administration
XX of a compound that inhibits the interaction of a sex steroid receptor
XX with a sex steroid response element of the Cyr61 (cysteine rich heparin-
XX binding protein) promoter. Cyr61 has cytostatic activity. An antibody (I)
XX which neutralises Cyr61 can be used to prevent or inhibit breast cancer

CC cell proliferation by blocking sex steroid induced and growth factor
CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal,
CC heparin binding epidermal or basic fibroblastic growth factor. (I) can be
CC used to diagnose or stage breast cancer where the level of Cyr61 in a
CC positive/suspect breast cancer cell is compared to the level in a normal
CC cell, an increase in the level of Cyr61 compared to the level in a normal
CC tissue indicates the presence of breast cancer. The level of Cyr61 being
CC determined by exposing the tissues to (I), and an increase in the level
CC of bound antibody by the suspect/positive cell as compared to the normal
CC tissue indicates the presence of breast cancer. The present sequence
CC represents the human Cyr61 protein, which is used in the exemplification
XX of the present invention
XX
SQ Sequence 381 AA;
Query Match 92.2%; Score 1938; DB 5; Length 381;
Best Local Similarity 91.4%; Pred. No. 1.9e-142;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
Qy 1 MSSSTFRTLAVALVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCGCKVCAKQL 60
Db 1 MSSRIARALAVVTLHLTRALSTCPAAACHPCLEAPKAPGVGLVRDGGCGCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENRIYONGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENRIYONGESFQPNCKHQ 120
Qy 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVSGQCEEWVCDSDIKSLDDDDLL-- 178
Db 121 CTCIDGAVGCIPLCPQELSLNLCNPNRLVKVSGQCEEWVCDSDIKSLDDDDLLG 180
Qy 179 --LGLDASEVELTRNNELIAGKSSLRKLPVFGTEPRVLPNPLHAHQKCIQVTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAGKSSLRKLPVFGTEPRVLPNPLHAHQKCIQVTTWSQ 238
Qy 237 CSKSGCTGISTRVTNDNPECLVKEVTRICEVRPCQPVYSSLRKGLKSKTKKSPPEVRF 296
Db 239 CSKSGCTGISTRVTNDNPECLVKEVTRICEVRPCQPVYSSLRKGLKSKTKKSPPEVRF 298
Qy 297 TYAGSSVKYRPRKYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNVMMIQSKCKNYN 356
Db 299 TYAGCLSVKYRPRKYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNVMMIQSKCKNYN 358
Qy 357 CPHPNEASFRLYSLFNDIHKFRD 379
Db 359 CPHPNEAAPPYRLFNDIHKFRD 381
RESULT 8
AAE18107
ID AAE18107 standard; protein; 381 AA.
XX
XX AAE18107;
XX
DT 07-MAY-2002 (first entry)
XX
XX Human connective tissue growth factor-2 (CTGF-2).
XX
XX Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
XX ischaemia; restenosis; tissue repair; wound healing; congenital defect;
XX cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
XX burns; osteoporosis; periodontal disease; liver failure; tranquilizer;
XX vulnery; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
XX gene therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Key Region 35..41
XX FT /note= "Immunogenic epitope"
XX FT Region 47..51
XX FT /note= "Immunogenic epitope"
XX FT Region 59..75

CC antibodies. The polynucleotide sequence encoding CTGF-2 may be used as a
 CC hybridisation probe for a cDNA library to isolate the full-length cDNA
 CC and other cDNAs, which have high sequence similarity to the CTGF-2 gene
 CC or similar biological activity. The polynucleotides are also useful for
 CC producing polypeptides by recombinant techniques, and in chromosome
 CC identification. The present sequence represents human CTGF-2. Note: The
 CC present sequence given as SEQ ID No:2 in the sequence listing differs
 CC from that given in Fig 1 (ABU63223)
 XX
 XX SQ Sequence 381 AA;
 Query Match 92.2%; Score 1938; DB 6; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.9e-142;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSSTFRLAVATLLHLTRALSTCPAACHCPLAPKAPGVGLVRDGGCCCKVCAKQL 60
 Db 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLAPKAPGVGLVRDGGCCCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGICPLCPQELSLPGLNLCNPRLVKVSGCCCEEWVCDSDSIKSLDDQDDL-- 178
 Db 121 CTCIDGAVGICPLCPQELSLPGLNLCNPRLVKVSGCCCEEWVCDSDSIKSLDDQDDL-- 178
 QY 179 --LGLDASEVELTRNELLTAIGKSLKRLPVFGTEPRVLPNPLHAGKQKCIQTTSWSQ 236
 Db 181 KELGFDASEVELTRNELLTAIGKSLKRLPVFGTEPRVLPNPLHAGKQKCIQTTSWSQ 238
 QY 237 CSKSGCTGISTRTVNDNPECLVKETRICVVRPCQPVYSSLLKGGKCKTKKSPPEVRF 296
 Db 239 CSKSGCTGISTRTVNDNPECLVKETRICVVRPCQPVYSSLLKGGKCKTKKSPPEVRF 298
 QY 297 TYAGCSVKKYRPKYCGSCVDGRCTCTPQTRVTKMRFCEDGEMFSKNVMMIQSKCNYN 356
 Db 299 TYAGCSVKKYRPKYCGSCVDGRCTCTPQTRVTKMRFCEDGEMFSKNVMMIQSKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 Db 359 CPHANEAAFPFYRLFNDIHKFRD 381
 RESULT 10
 AAB43987
 ID AAB43987 standard; protein; 455 AA.
 XX
 AC AAB43987;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1432.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW antidiabetic; antiaschmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antihypertensive; antiallergic; antidiabetic; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;
 KW vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disease; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO200005350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005882.
 XX

PR 12-MAR-1999; 99US-0124270P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX WPI; 2000-567533/55.
 DR N-PSDB; AAC78196.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 XX
 XX Claim 11; Page 2116-2118; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antiaschmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertensive; antiallergic; antidiabetic; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neutropenic; vasotrophic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 XX SQ Sequence 455 AA;
 Query Match 92.2%; Score 1938; DB 3; Length 455;
 Best Local Similarity 91.4%; Pred. No. 2.3e-142;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSSTFRLAVATLLHLTRALSTCPAACHCPLAPKAPGVGLVRDGGCCCKVCAKQL 60
 Db 75 MSSRIARALAVVTLHLTRALSTCPAACHCPLAPKAPGVGLVRDGGCCCKVCAKQL 134
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPNCKHQ 120
 Db 135 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYONGESFQPNCKHQ 194
 QY 121 CTCIDGAVGICPLCPQELSLPGLNLCNPRLVKVSGCCCEEWVCDSDSIKSLDDQDDL-- 178
 Db 195 CTCIDGAVGICPLCPQELSLPGLNLCNPRLVKVSGCCCEEWVCDSDSIKSLDDQDDL-- 254
 QY 179 --LGLDASEVELTRNELLTAIGKSLKRLPVFGTEPRVLPNPLHAGKQKCIQTTSWSQ 236
 Db 255 KELGFDASEVELTRNELLTAIGKSLKRLPVFGTEPRVLPNPLHAGKQKCIQTTSWSQ 312
 QY 237 CSKSGCTGISTRTVNDNPECLVKETRICVVRPCQPVYSSLLKGGKCKTKKSPPEVRF 296
 Db 313 CSKSGCTGISTRTVNDNPECLVKETRICVVRPCQPVYSSLLKGGKCKTKKSPPEVRF 372
 QY 297 TYAGCSVKKYRPKYCGSCVDGRCTCTPQTRVTKMRFCEDGEMFSKNVMMIQSKCNYN 356
 Db 373 TYAGCSVKKYRPKYCGSCVDGRCTCTPQTRVTKMRFCEDGEMFSKNVMMIQSKCNYN 432
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 Db 433 CPHANEAAFPFYRLFNDIHKFRD 455
 RESULT 11
 AAB76937

ID XX ABG76937 standard; protein; 381 AA.
AC XX ABG76937;
XX XX
DT 05-NOV-2002 (first entry)
DE XX
DE Human protein, comprising CYR61, designated SEC1.
XX XX
XX Human; SEC; NOV; immunosuppressive; hepatotropic; antiinflammatory;
KW angiogenic-associated disorder; diagnostic; gene therapy;
KW developmental disorder; immune disease; metabolic disorder;
KW signal transduction pathway disorder; neurodegenerative disorder;
KW feeding disorder; obesity; wasting disorder; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;
KW asthma; atherosclerosis; cardiomyopathy; angina pectoris;
KW autoimmune disease; retinal disease; cirrhosis; diabetes;
KW infectious disease; human immunodeficiency virus; HIV; cancer;
KW hypertension; hypotension; multiple sclerosis; urinary retention;
KW osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;
KW haemophilia; cirrhosis; immunogen; vaccine.
XX XX
OS Homo sapiens.
XX XX
XX WO200255705-A2.
XX XX
PD 18-JUL-2002.
XX XX
XX 11-JAN-2002; 2002WO-US000609.
XX XX
XX 11-JAN-2001; 2001US-0261013P.
XX 11-JAN-2001; 2001US-0261014P.
XX 11-JAN-2001; 2001US-0261018P.
XX 11-JAN-2001; 2001US-0261026P.
XX 11-JAN-2001; 2001US-0261029P.
XX 17-AUG-2001; 2001US-0313170P.
XX 10-SEP-2001; 2001US-0318410P.
XX XX
PA (CURA-) CURAGEN CORP.
XX XX
XX Mezes PS, Rastelli L, Hermann JL, MacDougall JR, Zhong H;
PI Casman SU, Boldog F, Shimkets RA, Gorman L, Crasta OR, Mysore KK;
PI Folkerts O, Martin GB, Eisen A, Spaderna SK, Vernet CAM, Bergh C;
PI Sytek KA, Dipippo VA, Zehusen BD, Peyman JA, Ellerman K, Stone DJ;
PI Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CB;
PI Edinger S;
XX XX
XX WPI; 2002-590675/63.
XX N-PSDB; ABS59522.
XX XX
XX Human SECX/NOVX polypeptide useful for diagnosing, preventing or treating
PT disorders associated with aberrant expression or activity of SECX/NOVX
PT nucleic acids and proteins e.g., diabetes.
XX XX
XX Example 3; Page 9; 443pp; English.
XX XX
XX The invention discloses the isolated human polypeptides, and
XX polynucleotides encoding them, that have been designated SECX and NOVX.
XX The polypeptides can be used for treating, or delaying, the onset of an
XX angiogenic-associated disorder or treating a pathological state in a
XX subject, preferably a mammal. They can also be used in determining the
XX presence of, or predisposition to, a disease associated with altered
XX levels of the polypeptides and polynucleotides of any one of the 12
XX sequences (SEC1-12), for raising antibodies, for identifying an agent
XX that binds to, or that modulates the expression or activity of the
XX polypeptide, for treating or preventing a NOVX-associated disorder (NOV1-
XX 8) and as a pharmaceutical composition comprising the polypeptide,
XX polynucleotide or the antibody. The polypeptides and polynucleotides are
XX useful in diagnostic applications where their amounts are assessed, or
XX for the manufacture of a medicament (e.g. gene therapy) for treating or
XX preventing disorders or syndromes such as developmental disorders, immune
XX diseases, signal transduction pathway disorders, metabolic disorders,
XX feeding disorders (including obesity), wasting disorders,
XX neurodegenerative disorders (including Alzheimer's disease and

CC Parkinson's disease), behavioural disorders, allergies, asthma,
CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They may
CC also be used as immunogens to produce antibodies specific for the
CC invention, and as vaccines. Further, they are useful for screening in
CC potential agonist and antagonist compounds. The sequences presented in
CC ABG76937-ABG76956 are the human SEC1-12 and NOV1-8 proteins
XX
XX Sequence 381 AA;
SQ
Query Match 91.8%; Score 1930; DB 5; Length 381;
Best Local Similarity 90.9%; Pred. No. 7.9e-142;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
Qy 1 MSSSTRFLAVAVTLLHLTLALSTCPAACHCELEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
Db 1 MSSRIAPALAVVTLHLTLALSTCPAACHCELEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
Qy 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRIVYQNGESFPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRIVYQNGESFPNCKHQ 120
Qy 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGCCCEWVCDSDSIKSLDDQDL-- 178
Db 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGCCCEWVCDSDSIKSLDDQDLG 180
Qy 179 --LGLDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIQVTTWSQ 236
Db 181 KELGPDASEVELTRNNELIAVGGSSLKRLPVFGTEPRVLFNPL--QGKCIQVTTWSQ 238
Qy 237 CSKSCGTGISTRVNDNPECLVKETRICVRCPCGPVYSSLLKGGKCKTKKSPVPVRF 296
Db 239 CSKTCGTGISTRVNDNPECLVKETRICVRCPCGPVYSSLLKGGKCKTKKSPVPVRF 298
Qy 297 TYAGCSVKYRKYKYGSCVDGRCCTPLOTTRTVMKFRCEGDSMFSGKNVMIOSCKCNYN 356
Db 299 TYAGCLSVKYRKYKYGSCVDGRCCTPLOTTRTVMKFRCEGDSMFSGKNVMIOSCKCNYN 358
Qy 357 CPHNPEASFLYSLFNDIHKFRD 379
Db 359 CPHANEAAFPYRLFNDIHKFRD 381
RESULT 12
ADB75260
ID ADB75260 standard; protein; 381 AA.
XX
AC ADB75260;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker protein.
XX
XX Prostate; cancer; cytostatic; gene therapy; marker.
XX
XX Homo sapiens.
XX
XX WO2003009814-A2.
XX
XX 06-FEB-2003.
XX
XX 25-JUL-2002; 2002WO-US023913.
XX
XX 25-JUL-2001; 2001US-0307982P.
XX 22-AUG-2001; 2001US-0314356P.
XX 25-SEP-2001; 2001US-0325020P.
XX 12-DEC-2001; 2001US-0341746P.
XX 05-MAR-2002; 2002US-0362158P.
XX XX

PA (MILL-) MILLENNIUM PHARM INC.
 XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
 PI Hoersht S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
 XX WPI; 2003-248033/24.
 XX
 XX New nucleic acid molecule, useful for diagnosing or treating prostate
 PT cancer.
 XX
 XX Disclosure; SEQ ID NO 84; 99pp; English.
 XX
 XX The invention relates to newly discovered cancer markers associated with
 CC the cancerous state of prostate cells. Also disclosed is a method of
 CC assessing whether a patient is afflicted with prostate cancer. The method
 CC of the invention involves assessing whether a patient is afflicted with
 CC prostate cancer by comparing the level of expression of a marker in a
 CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB/5177-ADB/5631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 381 AA;
 SQ
 Query Match 91.8%; Score 1930; DB 7; Length 381;
 Best Local Similarity 90.9%; Pred. No. 7.9e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
 QY 1 MSSSTRTTAVATLLHLTRALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRYQNGESFPQNCXHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRYQNGESFPQNCXHQ 120
 QY 121 CTCIDGAVGICPLCPQELSPLNLCNPNRLVKVSGCCBEWYCDSDSIKSLDDODDL-- 178
 DB 121 CTCIDGAVGICPLCPQELSPLNLCNPNRLVKVSGCCBEWYCDSDSIKSLDDODDL 180
 QY 179 --LGLDASEVELTRNNELIAVGKSSLKRLPVFGMEPRIRYNPL--QGQKCIQVTTWSQ 236
 DB 181 KELGFDASEVELTRNNELIAVGKSSLKRLPVFGMEPRIRYNPL--QGQKCIQVTTWSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKGGKCKSKTKSPVPVF 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKGGKCKSKTKSPVPVF 298
 QY 297 TYAGCSSVKYRPKYCGSCVDGRCCCTPLOTTRTVKMFRCEDGEMFKNVMMIOCKCNYN 356
 DB 299 TYAGCLSVKXKRPKYCGSCVDGRCCCTPLOTTRTVKMFRCEDGEMFKNVMMIOCKCNYN 358
 QY 357 CPHNEASFLYSLFNDIHKFRD 379
 DB 359 CPHANEAPFPYRLFNDIHKFRD 381
 RESULT 13
 AAW35730
 ID AAW35730 standard; protein; 381 AA.
 XX
 AC AAW35730;
 DT 27-MAR-1998 (first entry)
 XX Human cysteine rich protein 61 (Cyr61).

KW Cysteine rich protein 61; Cyr61; human;
 KW extracellular matrix signalling molecule; cell adhesion; cell migration;
 KW cell proliferation; angiogenesis; chondrogenesis; oncogenesis;
 KW haematostasis; wound healing; organ regeneration.
 OS Homo sapiens.
 XX
 XX W09733995-A2.
 XX 18-SEP-1997.
 XX
 XX 14-MAR-1997; 97WO-US004193.
 XX
 XX 15-MAR-1996; 96US-0013958P.
 XX (MUNI-) MUNIN CORP.
 XX Lau LF;
 XX WPI; 1997-470875/43.
 XX N-PSDB; AAT94699.
 XX Isolated and purified cysteine rich protein 61, Cyr61 - useful to
 PT modulate e.g. haematostasis, induce wound healing, promote organ
 PT regeneration etc.
 XX
 XX Claim 2; Page 112-113; 133pp; English.
 XX
 XX This protein sequence comprises human cysteine rich protein 61 (Cyr61),
 CC an extracellular matrix signalling molecule. Its amino acid sequence was
 CC deduced from a human placental cDNA clone (see AAT94699). Cyr61
 CC polypeptides can be expressed in transformed or transfected host cells.
 CC Cyr61 can be used to modulate haematostasis, induce wound healing in a
 CC tissue, promote organ regeneration, improve tissue grafting or promote
 CC bone or prothesis implantation (claimed). It can also be used to screen
 CC for a modulator of angiogenesis, chondrogenesis, oncogenesis, cell
 CC adhesion, cell migration, cell proliferation, expand a population of
 CC undifferentiated haematopoietic stem cells in culture and to screen for a
 CC mitogen (claimed). Ex vivo methods for using mammalian extracellular
 CC matrix signalling molecules to prepare blood products are also provided
 XX
 XX Sequence 381 AA;
 SQ
 Query Match 91.7%; Score 1929; DB 2; Length 381;
 Best Local Similarity 90.9%; Pred. No. 9.4e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
 QY 1 MSSSTRTTAVATLLHLTRALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRYQNGESFPQNCXHQ 120
 DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRYQNGESFPQNCXHQ 120
 QY 121 CTCIDGAVGICPLCPQELSPLNLCNPNRLVKVSGCCBEWYCDSDSIKSLDDODDL-- 178
 DB 121 CTCIDGAVGICPLCPQELSPLNLCNPNRLVKVSGCCBEWYCDSDSIKSLDDODDL 180
 QY 179 --LGLDASEVELTRNNELIAVGKSSLKRLPVFGMEPRIRYNPL--QGQKCIQVTTWSQ 236
 DB 181 KELGFDASEVELTRNNELIAVGKSSLKRLPVFGMEPRIRYNPL--QGQKCIQVTTWSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKGGKCKSKTKSPVPVF 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLKGGKCKSKTKSPVPVF 298
 QY 297 TYAGCSSVKYRPKYCGSCVDGRCCCTPLOTTRTVKMFRCEDGEMFKNVMMIOCKCNYN 356
 DB 299 TYAGCLSVKXKRPKYCGSCVDGRCCCTPLOTTRTVKMFRCEDGEMFKNVMMIOCKCNYN 358
 QY 357 CPHNEASFLYSLFNDIHKFRD 379

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

RESULT 1
 US-09-853-625B-11
 ; Sequence 11, Application US/09853625B
 ; Patent No. US20020049304A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN
 ; CECHTI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/853,625B
 ; FILING DATE: 14-May-2001
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/053,587
 ; FILING DATE: <unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MULLINS, J.G.
 ; REGISTRATION NUMBER: 33,073
 ; REFERENCE/DOCKET NUMBER: 325800-442
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 AMINO ACIDS

[illegible]

SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: "Mouse Cys61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-053-753-2

Query Match 100.0%; Score 2103; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.8e-166;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTRTTAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
DB 1 MSSSTRTTAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCENSRVYONGSFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCENSRVYONGSFQPNCKHQ 120
QY 121 CTICIDGAVGCIPLCPQELSLPGLGCPNPRLVKVSQCCEWVCDSDSIKSLDDQDDLLG 180
DB 121 CTICIDGAVGCIPLCPQELSLPGLGCPNPRLVKVSQCCEWVCDSDSIKSLDDQDDLLG 180
QY 181 LDASEVELTNNELIAGKSSILKRLPVFGTEPRVLFNPLHAHQKCIQVTTSSWQCSKS 240
DB 181 LDASEVELTNNELIAGKSSILKRLPVFGTEPRVLFNPLHAHQKCIQVTTSSWQCSKS 240
QY 241 CGTGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKGGKCKTKKSPVPVFTYAG 300
DB 241 CGTGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKGGKCKTKKSPVPVFTYAG 300
QY 301 CSSVKYRKYKSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
DB 301 CSSVKYRKYKSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
QY 361 NEASFRLYSIFNDIHKFRD 379
DB 361 NEASFRLYSIFNDIHKFRD 379

RESULT 2
US-10-053-753-2
Sequence 2, Application US/10053753
Publication No. US20020150986A1
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053,753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: "Mouse Cys61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-053-753-2

Query Match 100.0%; Score 2103; DB 13; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.8e-166;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTRTTAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
DB 1 MSSSTRTTAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRGCGCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCENSRVYONGSFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKICRAQSEGRPCENSRVYONGSFQPNCKHQ 120
QY 121 CTICIDGAVGCIPLCPQELSLPGLGCPNPRLVKVSQCCEWVCDSDSIKSLDDQDDLLG 180
DB 121 CTICIDGAVGCIPLCPQELSLPGLGCPNPRLVKVSQCCEWVCDSDSIKSLDDQDDLLG 180
QY 181 LDASEVELTNNELIAGKSSILKRLPVFGTEPRVLFNPLHAHQKCIQVTTSSWQCSKS 240
DB 181 LDASEVELTNNELIAGKSSILKRLPVFGTEPRVLFNPLHAHQKCIQVTTSSWQCSKS 240
QY 241 CGTGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKGGKCKTKKSPVPVFTYAG 300
DB 241 CGTGISTRVNDNPECLVKETRICVRPCGQPVYSSLLKGGKCKTKKSPVPVFTYAG 300
QY 301 CSSVKYRKYKSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
DB 301 CSSVKYRKYKSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
QY 361 NEASFRLYSIFNDIHKFRD 379
DB 361 NEASFRLYSIFNDIHKFRD 379

RESULT 3
US-10-099-322-45
Sequence 45, Application US/10099322
Publication No. US20030215449A1
GENERAL INFORMATION:
APPLICANT: Mezes et al.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-240CIP
CURRENT APPLICATION NUMBER: US/10/099,322
CURRENT FILING DATE: 2002-09-11
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/261,014
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/261,018
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/318,410
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/261,013
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/261,026
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/261,029
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: 60/313,170
PRIOR FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: 10/044,564
PRIOR FILING DATE: 2002-01-11
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45

```
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-099-322-45

Query Match
Best Local Similarity 100.0%; Score 2103; DB 15; Length 379;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
DB 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
QY 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
DB 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
QY 301 CSSVKYKRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
DB 301 CSSVKYKRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
QY 361 NEASPRLYSLFNDIHKFRD 379
DB 361 NEASPRLYSLFNDIHKFRD 379
```

```
RESULT 4
US-10-182-432-2
; Sequence 2; Application US/10182432
; Publication No. US200400021241
; GENERAL INFORMATION:
; APPLICANT: LAU, Lester F., YEUNG, Cho-Yau, and GREENSPAN, Jeffrey A.
; TITLE OF INVENTION: CYR61 COMPOSITIONS AND METHODS
; FILE REFERENCE: 214448/00029
; CURRENT APPLICATION NUMBER: US/10/182,432
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentih Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-182-432-2
```

```
Query Match
Best Local Similarity 100.0%; Score 2103; DB 15; Length 379;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
```

```
DB 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
QY 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
DB 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
QY 301 CSSVKYKRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
DB 301 CSSVKYKRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
QY 361 NEASPRLYSLFNDIHKFRD 379
DB 361 NEASPRLYSLFNDIHKFRD 379
```

```
RESULT 5
US-10-044-564-45
; Sequence 45; Application US/10044564
; Publication No. US200400018196A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240
; CURRENT APPLICATION NUMBER: US/10/044,564
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: Patentih Ver. 2.1
; SEQ ID NO 45
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-044-564-45
```

```
Query Match
Best Local Similarity 100.0%; Score 2103; DB 15; Length 379;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
DB 1 MSSSTFRTLAVAVTLHLTLRLALSTCPAACHCPCLEAPKCAPGVGLVRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCENSRINYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
DB 121 CTCIDGAVGCIPLCPQELSPLNLCNPNRLVKVSGQCCCEEWVCDSDSKSLDDQDQLLG 180
QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
DB 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHQKCIIVQTTWSQCSKS 240
QY 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
DB 241 CGTGISTRTVNDNPECLVKETRICVPRPCGQPVYSSLLKGGKCKSKTKKSPVPVRYTYAG 300
QY 301 CSSVKYKRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEGMFSKNVMMIOSCKNVCNCPHP 360
```

Db 301 CSSVKKYRKYCGSCVDGRCTPLQRTVQWPRCEDGEMFKNVMMIOCKNKNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 6
 US-10-464-368-60
 ; Sequence 60, Application US/10464368
 ; Publication No. US20040023356A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krümlauf, Robb
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
 ; FILE REFERENCE: 40716-IP-017
 ; CURRENT FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: US/10/464,368
 ; PRIOR FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: 60/388,970
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 60
 ; TYPE: PRT
 ; ORGANISM: MOUSE
 ; US-10-464-368-60

Query Match 100.0%; Score 2103; DB 16; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1.8e-166;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSSTRTTAVAVTLHLTLALSTCPAACHCPLEAPKCAPGVGLVRDGGCKVCCKQL 60
 Db 1 MSSSTRTTAVAVTLHLTLALSTCPAACHCPLEAPKCAPGVGLVRDGGCKVCCKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRIYQNGESFPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRIYQNGESFPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIAGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 180
 Db 121 CTCIDGAVGCIPLCPQELSLPNIAGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 180
 QY 181 LDASEVELTRNNELIATGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 240
 Db 181 LDASEVELTRNNELIATGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 240
 QY 241 CGTGISTRVNDNPECLRVKTRICEVRPCGQPVYSSLKGGKCKTKKSPPEVFTYAG 300
 Db 241 CGTGISTRVNDNPECLRVKTRICEVRPCGQPVYSSLKGGKCKTKKSPPEVFTYAG 300
 QY 301 CSSVKKYRKYCGSCVDGRCTPLQRTVQWPRCEDGEMFKNVMMIOCKNKNCPHP 360
 Db 301 CSSVKKYRKYCGSCVDGRCTPLQRTVQWPRCEDGEMFKNVMMIOCKNKNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 7
 US-10-464-368-62
 ; Sequence 62, Application US/10464368
 ; Publication No. US20040023356A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krümlauf, Robb
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
 ; FILE REFERENCE: 40716-IP-017
 ; CURRENT FILING DATE: 2003-06-16

; PRIOR APPLICATION NUMBER: 60/388,970
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 62
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: RAT
 ; US-10-464-368-62

Query Match 98.7%; Score 2076; DB 16; Length 379;
 Best Local Similarity 98.2%; Pred. No. 3.2e-164;
 Matches 372; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSSSTRTTAVAVTLHLTLALSTCPAACHCPLEAPKCAPGVGLVRDGGCKVCCKQL 60
 Db 1 MSSSTRTTAVAVTLHLTLALSTCPAACHCPLEAPKCAPGVGLVRDGGCKVCCKQL 60
 QY 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRIYQNGESFPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCYNRSRIYQNGESFPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIAGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 180
 Db 121 CTCIDGAVGCIPLCPQELSLPNIAGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 180
 QY 181 LDASEVELTRNNELIATGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 240
 Db 181 LDASEVELTRNNELIATGKSSLKRLPVGTEPRVLFNPLHAGQKCIQVTTTWSQCSKS 240
 QY 241 CGTGISTRVNDNPECLRVKTRICEVRPCGQPVYSSLKGGKCKTKKSPPEVFTYAG 300
 Db 241 CGTGISTRVNDNPECLRVKTRICEVRPCGQPVYSSLKGGKCKTKKSPPEVFTYAG 300
 QY 301 CSSVKKYRKYCGSCVDGRCTPLQRTVQWPRCEDGEMFKNVMMIOCKNKNCPHP 360
 Db 301 CSSVKKYRKYCGSCVDGRCTPLQRTVQWPRCEDGEMFKNVMMIOCKNKNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 8
 US-10-099-322-44
 ; Sequence 44, Application US/10099322
 ; Publication No. US20030215449A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes et al.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-240CIP
 ; CURRENT APPLICATION NUMBER: US/10/099,322
 ; CURRENT FILING DATE: 2002-09-11
 ; PRIOR APPLICATION NUMBER: 60/261,014
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,018
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/318,410
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/261,013
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,026
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,029
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/313,170
 ; PRIOR FILING DATE: 2001-08-17
 ; PRIOR APPLICATION NUMBER: 10/044,564
 ; PRIOR FILING DATE: 2002-01-11
 ; NUMBER OF SEQ ID NOS: 324
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 44
 ; LENGTH: 361

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-44

Query Match      92.3%; Score 1942; DB 15; Length 381;
Best Local Similarity 91.6%; Pred. No. 4.4e-153;
Matches 351; Conservative 9; Mismatches 17; Indels 6; Gaps 2;

QY 1 MSSSTRTTAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178

QY 179 --LGLDASEVELTRNNELIAIGKSSSLKRLPVFGTEPRVLFNPLHAHQKCIQVTTWSQ 236
DB 181 KELGFDASEVELTRNNELIAIGKSSSLKRLPVFGMEPRILYNPL--QGQKCIQVTTWSQ 238

QY 237 CSKSGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 296
DB 239 CSKTGCTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 298

QY 297 TYAGCSSVKYRKYKPGYCGSCVDGRCCTPQLOTRTVKMRFRCEDEGEMPSKNVMIQSCCKNYN 356
DB 299 TYAGCLSVKYRKYKPGYCGSCVDGRCCTPQLOTRTVKMRFRCEDEGEMPSKNVMIQSCCKNYN 358

QY 357 CPHNEASRLYSLFNDIHKFRD 379
DB 359 CPHANEAAFLYRLFNDIHKFRD 381

RESULT 9
US-10-044-564-44
; Sequence 44, Application US/10044564
; Publication No. US20040018196A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240
; CURRENT APPLICATION NUMBER: US/10/044,564
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-044-564-44

Query Match      92.3%; Score 1942; DB 15; Length 381;
Best Local Similarity 91.6%; Pred. No. 4.4e-153;
Matches 351; Conservative 9; Mismatches 17; Indels 6; Gaps 2;

QY 1 MSSSTRTTAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178

QY 179 --LGLDASEVELTRNNELIAIGKSSSLKRLPVFGTEPRVLFNPLHAHQKCIQVTTWSQ 236
DB 181 KELGFDASEVELTRNNELIAIGKSSSLKRLPVFGMEPRILYNPL--QGQKCIQVTTWSQ 238

QY 237 CSKSGTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 296
DB 239 CSKTGCTGISTRTVNDNPECLRVKTRICEVRPCGQPVYSSLLKGGKCKSKTKKSPVPVR 298

QY 297 TYAGCSSVKYRKYKPGYCGSCVDGRCCTPQLOTRTVKMRFRCEDEGEMPSKNVMIQSCCKNYN 356
DB 299 TYAGCLSVKYRKYKPGYCGSCVDGRCCTPQLOTRTVKMRFRCEDEGEMPSKNVMIQSCCKNYN 358

QY 357 CPHNEASRLYSLFNDIHKFRD 379
DB 359 CPHANEAAFLYRLFNDIHKFRD 381

RESULT 9
US-09-901-910-2
; Sequence 2, Application US/09901910
; Publication No. US20030012768A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark
; APPLICANT: Calenda, Valerie
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PFI26P2
; CURRENT APPLICATION NUMBER: US/09/901,910
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; PRIOR APPLICATION NUMBER: 60/217,402
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/291,642
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-901-910-2

Query Match      92.2%; Score 1938; DB 10; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTRTTAVAVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVVTLLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKVSGCCCEWVCDSDSIKDSLDDODDL-- 178

```


Db 121 CTCIDGAVGCIPLCPOELSLPNCNPRLVKVTGCCCEWYCDSDSIKDPMEDQDGLG 180
QY 179 --LGLDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLPNPLHAHOKCIVQTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAVKGSSLLKRLPVFGMEPRILYNPL--QOQKIVQTTWSQ 238
QY 237 CSKSCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 296
Db 239 CSKTCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 298
QY 297 TVAGSSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 356
Db 299 TVAGCLSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 358
QY 357 CPHNEASPRLYSLFNDIHKFRD 379
Db 359 CPHANEAPFPYRLFNDIHKFRD 381

RESULT 11
US-10-294-796-2
; Sequence 2, Application US/10294796
; Publication No. US20030078391A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong et al.
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; CURRENT APPLICATION NUMBER: US/10/294,796
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-796-2

Query Match 92.2%; Score 1938; DB 14; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTPRTLAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
Db 1 MSSRTARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCENSRITYONGESFPQNCXQ 120
Db 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCENSRITYONGESFPQNCXQ 120
QY 121 CTCIDGAVGCIPLCPOELSLPNCNPRLVKVTGCCCEWYCDSDSIKDLDDDDL-- 178
Db 121 CTCIDGAVGCIPLCPOELSLPNCNPRLVKVTGCCCEWYCDSDSIKDPMEDQDGLG 180
QY 179 --LGLDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLPNPLHAHOKCIVQTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAVKGSSLLKRLPVFGMEPRILYNPL--QOQKIVQTTWSQ 238
QY 237 CSKSCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 296
Db 239 CSKTCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 298
QY 297 TVAGSSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 356
Db 299 TVAGCLSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 358
QY 357 CPHNEASPRLYSLFNDIHKFRD 379
Db 359 CPHANEAPFPYRLFNDIHKFRD 381

Db 359 CPHANEAPFPYRLFNDIHKFRD 381

RESULT 12
US-10-394-015-5
; Sequence 5, Application US/10394015
; Publication No. US20030180891A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Connective Tissue Growth Factor-4
; FILE REFERENCE: P467
; CURRENT APPLICATION NUMBER: US/10/394,015
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US/09/325,019
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/088,320
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-015-5

Query Match 92.2%; Score 1938; DB 14; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTPRTLAVAVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
Db 1 MSSRTARALAVVTLHLTRLALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKQL 60
QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCENSRITYONGESFPQNCXQ 120
Db 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPCENSRITYONGESFPQNCXQ 120
QY 121 CTCIDGAVGCIPLCPOELSLPNCNPRLVKVTGCCCEWYCDSDSIKDLDDDDL-- 178
Db 121 CTCIDGAVGCIPLCPOELSLPNCNPRLVKVTGCCCEWYCDSDSIKDPMEDQDGLG 180
QY 179 --LGLDASEVELTRNNELIAIGKSSLLKRLPVFGTEPRVLPNPLHAHOKCIVQTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAVKGSSLLKRLPVFGMEPRILYNPL--QOQKIVQTTWSQ 238
QY 237 CSKSCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 296
Db 239 CSKTCGTGISTRTVNDNPECLVKTETRICVAPPCGQPVYSSLLKGGCKSKTKKSPVPVF 298
QY 297 TVAGSSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 356
Db 299 TVAGCLSVKKYRPKYCGSCVDCRCCTPLQTRTVKMRFCEDGEMFSKNVMIQSKCKNYN 358
QY 357 CPHNEASPRLYSLFNDIHKFRD 379
Db 359 CPHANEAPFPYRLFNDIHKFRD 381

RESULT 13
US-10-099-322-42
; Sequence 42, Application US/10099322
; Publication No. US20030215449A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/10/099,322
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11

```
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-42

Query Match          92.2%; Score 1938; DB 15; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFRTLAVAVTLLHLTRLALSTCPAAACHCPLKAPGVLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALALVVTLHLTRLALSTCPAAACHCPLKAPGVLVRDGGCCCKVCAKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCVNSRIYONGESFPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCVNSRIYONGESFPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPOELSLPMLGCPNRLVKVSGCCCEWVCDSDSIKSLDDQDDL-- 178
DB 121 CTCIDGAVGCIPLCPOELSLPMLGCPNRLVKVSGCCCEWVCDSDSIKSLDDQDDL 180

QY 179 --LGLDASEVELTRNNELIATGKSSSLKRLPVFGTEPRVLNPLHAHOKKCIIVOTTSWQ 236
DB 181 KELGPDASEVELTRNNELIATGKSSSLKRLPVFGTEPRVLNPLHAHOKKCIIVOTTSWQ 238

QY 237 CSKSCGTGISTRVNDNPECLVKEITRICEVRPCQOPVYSSLKKGKCKSKTKSPPEVRF 296
DB 239 CSKTCGTGISTRVNDNPECLVKEITRICEVRPCQOPVYSSLKKGKCKSKTKSPPEVRF 298

QY 297 TYAGCSVKKYRPKYCGSCVDGRCTPLOTRTVYKRFCEDEGEMESKNVMIQSKCNYN 356
DB 299 TYAGCLSVKKYRPKYCGSCVDGRCTPLOTRTVYKRFCEDEGEMESKNVMIQSKCNYN 358

QY 357 CPHNEAGFRLYSLFNDIHKFRD 379
DB 359 CPHANEAFFPYRLFNDIHKFRD 381

RESULT 14
US-10-044-564-42
; Sequence 42, Application US/10044564
; Publication No. US20040018196A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240
; CURRENT APPLICATION NUMBER: US/10/044,564
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11

; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-322-42

Query Match          92.2%; Score 1938; DB 15; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFRTLAVAVTLLHLTRLALSTCPAAACHCPLKAPGVLVRDGGCCCKVCAKQL 60
DB 1 MSSRIARALALVVTLHLTRLALSTCPAAACHCPLKAPGVLVRDGGCCCKVCAKQL 60

QY 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCVNSRIYONGESFPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLCNFGASSTALGICRAQSEGRPCVNSRIYONGESFPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPOELSLPMLGCPNRLVKVSGCCCEWVCDSDSIKSLDDQDDL-- 178
DB 121 CTCIDGAVGCIPLCPOELSLPMLGCPNRLVKVSGCCCEWVCDSDSIKSLDDQDDL 180

QY 179 --LGLDASEVELTRNNELIATGKSSSLKRLPVFGTEPRVLNPLHAHOKKCIIVOTTSWQ 236
DB 181 KELGPDASEVELTRNNELIATGKSSSLKRLPVFGTEPRVLNPLHAHOKKCIIVOTTSWQ 238

QY 237 CSKSCGTGISTRVNDNPECLVKEITRICEVRPCQOPVYSSLKKGKCKSKTKSPPEVRF 296
DB 239 CSKTCGTGISTRVNDNPECLVKEITRICEVRPCQOPVYSSLKKGKCKSKTKSPPEVRF 298

QY 297 TYAGCSVKKYRPKYCGSCVDGRCTPLOTRTVYKRFCEDEGEMESKNVMIQSKCNYN 356
DB 299 TYAGCLSVKKYRPKYCGSCVDGRCTPLOTRTVYKRFCEDEGEMESKNVMIQSKCNYN 358

QY 357 CPHNEAGFRLYSLFNDIHKFRD 379
DB 359 CPHANEAFFPYRLFNDIHKFRD 381

RESULT 15
US-10-381-644-2
; Sequence 2, Application US/10381644
; Publication No. US20040023910A1
; GENERAL INFORMATION:
; APPLICANT: American Home Products Corporation
; APPLICANT: Zhang, Zhiming
; APPLICANT: Sampath, Deepak
; APPLICANT: Zhu, Yuan
; APPLICANT: Winnekar, Richard
; TITLE OF INVENTION: Use of Cyt61 in the treatment and
; FILE REFERENCE: AM100352
; CURRENT APPLICATION NUMBER: US/10/381,644
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/236,887
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-381-644-2

Query Match          92.2%; Score 1938; DB 16; Length 381;
Best Local Similarity 91.4%; Pred. No. 9.5e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
```

```

QY 1 MSSSTFRTIAVATLHLTRLALSTCPAAACHCPLEAPKCAPGVGLVRDGGCGCKVCAKOL 60
Db 1 MSSRIARALAVVTLHLTRLALSTCPAAACHCPLEAPKCAPGVGLVRDGGCGCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLECNFGASSTALKICRAQSEGRECEYNSRIYONGESFQPNCKHQ 120
Db 61 NEDCSKTQPCDHTKGLECNFGASSTALKICRAQSEGRECEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVSGCCBEWYCDSDSINDSLDDODDL-- 178
Db 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVSGCCBEWYCDSDSINDSLDDODDL 180
QY 179 --LGLDASEVELTRNNELIAIKGSSIKELPVGTEPRVLFNPLHAHGOKCIVQTTWSQ 236
Db 181 KELGFDASEVELTRNNELIAVGKSSLKRLPVFGMEPRILYNPL--QGQKCIQVTTWSQ 238
QY 237 CSKSCGTGISTRVTNDNPECLVKETRICEVAPCGQPVYSSLLKGGKCKSKTKKSPVPVRF 296
Db 239 CSKTCGTGISTRVTNDNPECLVKETRICEVAPCGQPVYSSLLKGGKCKSKTKKSPVPVRF 298
QY 297 TYAGCSVVKYRPKYCGSCVDGRCCCTPLOTRTVYKRRFRCEDEMFSKNWMIOQCKKNYN 356
Db 299 TYAGCLSVVKYRPKYCGSCVDGRCCCTPQLTRTVYKRRFRCEDETFSKNWMIOQCKKNYN 358
QY 357 CPHNEASPRLYSLFNDIHKFRD 379
Db 359 CPHANEAPFFYLFNDIHKFRD 381

```

Search completed: April 22, 2004, 18:08:16
 Job time : 39.3987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:53:39 ; Search time 16.4566 Seconds
(without alignments)
2215.320 Million cell updates/sec

Title: US-09-495-448A-2

Perfect score: 2103

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFLYSLFNDIHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2103	100.0	379	2 A35669	Gene CYR61 protein
2	1645.5	78.3	375	2 A41428	CBF-10 protein pre
3	957.5	45.5	348	2 A40578	beta IG-M2 protein
4	950	45.2	349	2 A40551	connective tissue
5	861	40.9	351	2 S20078	NOV protein - chic
6	839.5	39.9	357	2 T38059	gene novH protein
7	171	8.1	1111	2 T26972	hypothetical prote
8	166	7.9	1034	2 J05598	mucin - rat
9	157.5	7.5	13288	2 T03099	mucin, submaxillar
10	156.5	7.4	1574	2 T13954	MEGF6 protein - ra
11	156.5	7.4	1620	2 T27283	hypothetical prote
12	156	7.4	1025	2 T42626	secreted leucine-r
13	152.5	7.3	1042	2 A57534	mucin 5AC (clone L
14	152	7.2	1056	2 A53767	mucin MUC5B, trach
15	151.5	7.2	1700	2 S08167	Balbani ring 3 pr
16	146	6.9	837	2 A42112	mucin-like peptide
17	145	6.9	601	2 T22025	hypothetical prote
18	145	6.9	601	2 D99711	protein P40810.4 (
19	145	6.9	1531	2 T42218	elit-1 protein hom
20	144.5	6.9	1170	2 A36612	laminin Blk chain
21	143	6.8	1964	2 T09059	notch4 - mouse
22	142.5	6.8	1101	2 T16840	hypothetical prote
23	141	6.7	251	2 A55035	cysteine-rich prot
24	141	6.7	5376	2 T42215	zonadhesin - mouse
25	140	6.7	3106	1 S53868	laminin alpha-2 ch
26	139.5	6.6	1178	1 A39804	thrombospondin pre
27	139	6.6	473	2 A56175	adhesive plaque pr
28	138.5	6.6	1847	2 T18308	probable vitellog
29	136.5	6.5	3020	2 A43932	mucin 2 precursor

RESULT 1

A35669

Gene CYR61 protein precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 28-Sep-1990 #sequence revision 18-Nov-1992 #text_change 05-Nov-1999

C;Accession: A35669; I48319; S16446

R;O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.

Mol. Cell. Biol. 10, 3569-3577, 1990

A;Title: Expression of cyr61, a growth factor-inducible immediate-early gene.

A;Reference number: A35669; M01D:90287146; F01D:2355916

A;Accession: A35669

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-379 <OAB>

A;Cross-references: GB:M32490; NID:g192909; PIDN:AAA37512.1; PID:g309206

A;Note: the authors translated the codon GAT for residue 337 as Gln

R;Jatinkic, B.V.; O'Brien, T.P.; Lau, L.F.

Nucleic Acids Res. 19, 3261-3267, 1991

A;Title: Promoter function and structure of the growth factor-inducible immediate early

A;Reference number: I48319; M01D:91288203; F01D:2062642

A;Accession: I48319

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-379 <RES>

A;Cross-references: EMBL:X56790; NID:g50632; PIDN:CAA40109.1; PID:g50633

A;Note: the authors did not translate the codon for residue 108

A;Note: the authors translated the codon GAT for residue 337 as Gln

C;Genetics:

A;Gene: CYR61

A;Introns: 21/3; 93/1; 208/1; 279/3

C;Superfamily: von Willebrand factor type C repeat homology

F;99-166/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match

Best Local Similarity 100.0%; Score 2103; DB 2; Length 379;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRTLAVALTLHLTRALSTCPAACHCPLPAPKAPGVGLVRGCGGCKVC AKQL 60

Db 1 MSSSTFRTLAVALTLHLTRALSTCPAACHCPLPAPKAPGVGLVRGCGGCKVC AKQL 60

QY 61 NEDCSKTQPCDHTKLECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFQNCXHQ 120

Db 61 NEDCSKTQPCDHTKLECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFQNCXHQ 120

QY 121 CTCIDGAVGCIFLCPQELSLPNLGNPCNRLVKVSCQCEEWVCDBDSIKSLDDDDLLIG 180

Db 121 CTCIDGAVGCIFLCPQELSLPNLGNPCNRLVKVSCQCEEWVCDBDSIKSLDDDDLLIG 180

QY 181 LDASEVELTNNELTAIGKSSLKELPVFGTEPRVLEPLHAHGOKCIVQTTSWSQCSKS 240

Db 181 LDASEVELTNNELTAIGKSSLKELPVFGTEPRVLEPLHAHGOKCIVQTTSWSQCSKS 240

QY 241 CGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVFTYAG 300
 Db 241 CGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVFTYAG 300
 QY 301 CSSVKYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCNCPHP 360
 Db 301 CSSVKYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCNCPHP 360
 QY 361 NEASFRLYSLENDIHKFRD 379
 Db 361 NEASFRLYSLENDIHKFRD 379

RESULT 2

A41428
 CEF-10 protein precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
 C:Accession: A41428
 R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Eriksen, R.L.
 Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
 A:Title: Identification of a phorbol ester-repressible v-src-inducible gene.
 A:Reference number: A41428; MUID:89145206; PMID:2537491
 A:Accession: A41428
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-375 <SIM>
 A:Cross-references: GB:J04496; NID:g211435; PIDN:AAA48661.1; PID:g211436

Query Match 78.3%; Score 1646.5; DB 2; Length 375;
 Best Local Similarity 80.1%; Pred. No. 3.1e-109;
 Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;
 QY 10 AVAVTLLHLTRLAL-STCPAAACHPCLEAPKCAPGVLYVDGCGCKKVCAGKQINEDCSKTQ 68
 Db 9 ALAAALLCLARLALGSPFAVCCPAAAPQCAPGVLYVDGCGCKKVCAGKQINEDCSKTQ 68
 QY 69 PCDHHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFPNCKHCTCIDGAV 128
 Db 69 PCDHHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFPNCKHCTCIDGAV 128
 QY 129 GCIPLCPOELSLNLCGNPRLVKVSGQCEWVDEDSIKSLDDDDLL-----LGLDAS 184
 Db 129 GCIPLCPOELSLNLCGNPRLVKVSGQCEWVDEDSIKSLDDDDLL-----LGLDAS 186
 QY 185 EVELTRNNELIATKGLSSLLKRLPVFGTEP--RVLFNPLHAHQKQICIVQTTSSQCSKSG 242
 Db 187 EVELTRNNELIATKGLSSLLKRLPVFGTEP--RVLFNPLHAHQKQICIVQTTSSQCSKSG 239
 QY 243 TGISSTRVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVFTYAGCS 302
 Db 240 TGISSTRVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVFTYAGCS 299
 QY 303 SVVKYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCNCPHPNE 362
 Db 300 SVVKYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCNCPHANE 359
 QY 363 ASFRLYSLENDIHKFRD 379
 Db 360 A-YPFVRLVNDIHKFRD 375

RESULT 3

A40578
 beta IG-M2 protein precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
 C:Accession: A40578; A53228
 R:Brunker, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
 DNA Cell Biol. 10, 293-300, 1991
 A:Title: Identification of a gene family regulated by transforming growth factor-beta.
 A:Reference number: A40578; MUID:91229699; PMID:2029937

A:Accession: A40578
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-348 <BRU>
 A:Cross-references: GB:M80263; NID:g201945; PIDN:AAA73135.1; PID:g201946
 R:Rybeck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
 Cell Growth Differ. 2, 225-233, 1991
 A:Title: Structure, mapping, and expression of fisp-12, a growth factor-inducible gene
 A:Reference number: A53228; MUID:91363290; PMID:1888698
 A:Accession: A53228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-160, 'K', 162-348 <RYS>
 A:Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
 C:Genetics:
 A:Gene: fisp-12

Query Match 45.5%; Score 957.5; DB 2; Length 348;
 Best Local Similarity 46.8%; Pred. No. 1.5e-60;
 Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;
 QY 1 MSSSTFTFLAVAVTLLHL-TRLAL-STCPAAACHPCLE-APKCAPGVLYVDGCGCKKVCVA 57
 Db 1 MLASVAGPISTALVLLALCTRPATGQCSAQCCAAEAAAPHCAPAGVSLVLDGCGCCRVCA 60
 QY 58 KQLNEDCSKTQPCDHHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQNGESFPNCK 117
 Db 61 KQLGELCTERDPCDPHKGFLCDFGSPANRKLIGVCTAK-DGAPCVFGSVYRSGESFSQSSC 119
 QY 118 KHQCTCIDGAVGCIPLCPQELSLNLCGNPRLVKVSGQCEWVDEDSIKSLDDDD 177
 Db 120 KYQCTCLDGVAGVCPVLCSDMVRVLPSPCPFPFRVRLFGKCCCEWVDEP----- 168
 QY 178 LLGLDASEVELTRNNELIATKGLSSLLKRL-PVFGTEPVLNPLHAHQKQICIVQTTSSQ 236
 Db 169 -----KRTAVGPALAAVRLDTGDPFTM-----RANCLVQTTEWSA 207
 QY 237 CSKSCGFGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVRF 296
 Db 208 CSKTCGNGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLLKGGKCKSKTKSPVPVRF 267
 QY 297 TYAGCSVKYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCN 356
 Db 268 ELSGCTSVKTYRKYCGSCVDGRCCTPLQTRTVKMRFRCEDEGMFSKNVMIIQSCCKNVCN 327
 QY 357 CPHNEASFRLY--SLFNDI 374
 Db 328 CPGDNDIFESLYRKYMGDM 347

RESULT 4

A40551
 connective tissue growth factor - human
 C:Species: Homo sapiens (man)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
 C:Accession: A40551; S44205
 R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
 J. Cell Biol. 114, 1285-1294, 1991
 A:Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human vas
 A:Reference number: A40551; MUID:91373462; PMID:1654338
 A:Accession: A40551
 A:Molecule type: mRNA
 A:Residues: 1-349 <BRA>
 A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AAA91279.1; PID:R:
 R:Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.; Luescher, T.F.
 submitted to the EMBL Data Library, April 1994
 A:Description: Differential cloning and expression of human connective tissue growth fa
 A:Reference number: S44205
 A:Accession: S44205
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-349 <OW>
 A:Cross-references: EMBL:X78947; NID:g474933; PID:g474934

Query Match 45.2%; Score 950; DB 2; Length 349;
Best Local Similarity 45.7%; Pred. No. 5e-60;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRTLAVAVTLHLTRALSTCPAA-----CHCPLE-APKCAPGVGLVRDGGC 52
DB 1 MTAASGPPVRVAFVVL-----LALCRPAVGQNCSPFCRCPDEPAPRCAGVSLVLDGGC 56

QY 53 KVKCAQLNEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCYNSRIYQGES 112
DB 57 CRVCAKQLGELCTERDPCDPKHGFLCDFGSPANRIGVCTAK-DGAPCIFGTVYRSES 115

QY 113 FQPNCKHCTCTDGVAGCIPLCQELSLPNLGCNPRLVKVSGQCCCEWVDEDSIKDSL 172
DB 116 FOSSCKYQCTLDGAVGCMPLCSMDVRUPSPDCPFPRRVKLPFGKCCCEWVDEPK----- 170

QY 173 DQDDLLGLDASEVELTRNNELIAIGKSSLRKL-PVFGTSPRVLNPLHAHQKCIYQT 231
DB 171 -DQ-----TVGPAALAAVRLDTFGDPMTI-----RANCLVQT 203

QY 232 TWSQCSKSCGIGISTRTVNDNPECLVKETRICVPRCGQPVYSLKKGKCKSKTKKSP 291
DB 204 TWSACSKTCGIGISTRTVNDNASCRLKQSLCVRPCADLEENIKKGGKCIKRTPKIS 263

QY 292 EPRVFTYAGCSSVKYRKYCGSCVDGRCTPLQTRTVKMRPRCEDGMPKSNVMMIQSC 351
DB 264 KPIKELSGCTSMKYRAKFCGVCCTDGRCTPHRTITLIPVEFKPDGVMKMMFIKTC 323

QY 352 KYNPCPNHNEASFRL--SLFNDI 374
DB 324 ACHYNCPGDNDFESLYRKMVGDM 348

RESULT 5
S20078
NOV protein - chicken
C/Species: Gallus gallus (chicken)
C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C/Accession: S20078
R/Joliot, V.; Martinerie, C.; Dambrine, G.; Plassiat, G.; Brisac, M.; Crochet, J.; Perth
Mol. Cell. Biol. 12, 10-21, 1992
A/Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myel
A/Reference number: S20078; MUID:92107157; PMID:1309586
A/Accession: S20078
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-351 <JOL>
A/Cross-references: EMBL:X59284; NID:G63702; PIDN:CAA41975.1; PID:G63703
C/Genetics:
A/Gene: NOV

Query Match 40.9%; Score 861; DB 2; Length 351;
Best Local Similarity 44.5%; Pred. No. 9.9e-54;
Matches 165; Conservative 44; Mismatches 110; Indels 52; Gaps 7;

QY 9 LAVAVTLHLTRLA-----LSTCPAC--HCPLKAPKAPGVGLVRDGGCCKKCAQLN 61
DB 9 LPVLLLLLLLRPCVSGREACRPPCGRCAPFPRCPAGVPAPVLDGCGCLVCARQG 68

QY 62 EDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCYNSRIYQGESFQPNCKHQ 121
DB 69 ESCSPLLPCEDESGGLYCDRGPEDEGGAGICMV-LEGDNCVFDGMVYRNGETFPQCKYQC 127

QY 122 TCIDGAVGICPLCPQELSLPNLGCNPRLVKVSGQCCCEWVCD-EDSI-----KDSLDDQ 175
DB 128 TCRDQIGICLPRCNLGLLPGDCPFPRKIEVPGECEKMKVCDPRDEVLGLGFAAAYRQ 187

QY 176 DDLGLDASEVELTRNNELIAIGKSSLRKLPLVFGTSPRVLNPLHAHQKCIYQTTWS 235
DB 198 EATLGDIVSD-----SSANCIEQTTEWS 210

QY 236 QCSKSCGTGISTRTVNDNPECLVKETRICVPRCGQPVYSLKKGKCKSKTKKSPBPVR 295

DB 211 ACSKSCMGFPSTRTVNRNQCEMVKQTRLCMMPECENEBPSD-KGKKCIQTKKMKAVR 269
QY 296 FTYAGCSSVKYRKYCGSCVDGRCTPLQTRTVKMRPRCEDGMPKSNVMMIQSCCKNY 355
DB 270 FEYANCSTVQTYAPRYICGLNDRGCTPHNTKTIQVEFRCPQKFLKKPMLLINTVCCHG 329

QY 356 NCPHPNEASFR 366
DB 330 NCPQSNNAFFQ 340

RESULT 6
I38069
gene novH protein - human
C/Species: Homo sapiens (man)
C/Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C/Accession: I38069
R/Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Perbal
Oncogene 9, 2729-2732, 1994
A/Title: Structural analysis of the human nov proto-oncogene and expression in Wilms tu
A/Reference number: I38069; MUID:94336229; PMID:7520150
A/Accession: I38069
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-357 <RES>
A/Cross-references: EMBL:X78351; NID:G587422; PIDN:CAA55146.1; PID:G825696
C/Genetics:
A/Gene: novH

Query Match 39.8%; Score 839.5; DB 2; Length 357;
Best Local Similarity 43.8%; Pred. No. 3.3e-52;
Matches 163; Conservative 52; Mismatches 116; Indels 41; Gaps 9;

QY 1 MSSSTF-----RTLAVAVTLHLTRALST--CPAAC--HCPLKAPKAPGVGLVRDGGC 52
DB 4 VQSTSFCLRKQCLCLFTLLHLGLQVAATQRCPPQCPGRCPTATPTTCAPGVRAVLDDGSC 63

QY 53 CKVCAQLNEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCYNSRIYQGES 112
DB 64 CLVCARQSGESCDLEPCDESGLYCDRSADPSNQTGICTA-VEGDNCFDGVYRSGBK 122

QY 113 FQPNCKHCTCTDGVAGCIPLCQELSLPNLGCNPRLVKVSGQCCCEWVDEDSIKDSL 172
DB 123 FQPSCKFQCTCRDGGQICVPRCQLDVLPEPNCAPRKVEVGECEKWCIPDE-EDSL 181

QY 173 DQDDLLGLDASEVELTRNNELIAIGKSSLRKLPLVFGTSPRVLNPLHAHQKCIYQTT 232
DB 182 GG-----LTAAVRPEATLGEV-----SDSSV-----NCIEQTT 211

QY 233 SWSQCSKSCGTGISTRTVNDNPECLVKETRICVPRCGQ-PVYSLKKGKCKSKTKKSP 291
DB 212 EWTACSKSCMGFSITVNRNQCCEMLKQTRLCWVRPCEQEPQPTDKGKKCLRTKSL 271

QY 292 EPRVFTYAGCSSVKYRKYCGSCVDGRCTPLQTRTVKMRPRCEDGMPKSNVMMIQSC 351
DB 272 KAIHLQFNKCTSLHTYKPRFCVCGSDGRCTPHNTKTIQAEFQCSGPGQIVKPKVMVIGTC 331

QY 352 KYNPCPNHNEA 363
DB 332 TCHTNCPRNEA 343

RESULT 7
T26972
hypothetical protein Y479C.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C/Accession: T26972
R/Harris, B.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20293
 A;Accession: J26972
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1111 <MIL>
 A;Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4
 A;Experimental source: clone Y47H9C
 C;Genetics:
 A;Gene: CESP:Y47H9C.4
 A;Map position: 1
 A;Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1
 C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match 8.1%; Score 171; DB 2; Length 1111;
 Best Local Similarity 20.1%; Pred. No. 0.00017;
 Matches 86; Conservative 32; Mismatches 156; Indels 154; Gaps 20;
 QY 26 CPAACHCPLE-APKAPGVGLR-----DGGCGKVCIAK-QLNEDCSKTQPCDHTKGLKC 78
 Db 380 CSKTCTCVRENTLMCAPTGFCRCKPGFYGDNCCLACSKDSYGPNCCKQAMCDWNHASEC 439
 QY 79 NFGASSTALGICRAQSEGRPCENSHIYQNGSFOPNCKHQCCTC-----IDGA----- 127
 Db 440 NPETGSC-----VCKPGRGKNCSEPCPL-----DFYGNCAHQCCNORGVGCDGADGKCQ 491
 QY 128 -----VGCIPL-----CPOELSLPN--LGCP-- 146
 Db 492 CDGWTGHRCEHHCPADTFGANCEKCKCPKIGICDPTTGECTCPAGLQGANCDIGCPRG 551
 QY 147 -----NPLRVKYGOCCEBHWDESDISKSLDDDDLLGLDASEVELTRNNE 193
 Db 552 SYPGCKLHKCVNGKDKETGEC---TC-----QGFPGSDCS----- 587
 QY 194 LIAIGKSSLLRLPFGTEPRVLNPLHAHQKCIQV--TTSWSQCSK---SGGTGISTRV 249
 Db 588 -TTCCKGK-----YGESCELSCPCSDASCSCSKQTKGKCLPLGTGK 625
 QY 250 TNDNPECLVETRICVPRCGOPVYSSLLKCKKCKTKSKPEPVFTY-----AGC 301
 Db 626 VSCDQKCDPNTFGLQETVTPFCASTDPKNGVCLSCPFGSGGIHCENCPAGSYGDCG 685
 QY 302 SSVKRYRPKYCGSDGRCCTPLQTRTV-----KMRPRCBDEGMFSKNVMMIQSKCK 353
 Db 686 QQV-----C-SCADHGCDPTTGECICPEGVHGTCEKCPDGGKYGVGCALDCPKCAS 737
 QY 354 NYNCPHN 361
 Db 738 GSTCDHIN 745

RESULT 8
 JC5598
 Mucin - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 05-Nov-1999
 C;Accession: JC5598
 R;Inatomi, T.; Tiedale, A.S.; Zhan, Q.; Spurr-Michaud, S.; Gipson, I.K.
 Biochem. Biophys. Res. Commun. 236, 789-797, 1997
 A;Title: Cloning of rat Muc5AC mucin gene: Comparison of its structure and tissue distribution
 A;Reference number: JC5598; MUID:97396181; PMID:9245735
 A;Accession: JC5598
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-1034 <INA>
 A;Cross-references: GB:U83139; NID:G2315984; PIDN:AA53312.1; PID:G2315985
 A;Note: translation not complete
 C;Comment: This protein is a high molecular weight glycoprotein which is a major component of intestinal tract and reproductive tract.
 A;Genetics:
 A;Gene: Muc5A
 C;Superfamily: von Willebrand factor type C repeat homology
 F;45-149/Domain: cysteine-rich <CS>

F;762-830/Domain: von Willebrand factor type C repeat homology <WC>

Query Match 7.9%; Score 166; DB 2; Length 1034;
 Best Local Similarity 21.4%; Pred. No. 0.00035;
 Matches 99; Conservative 47; Mismatches 170; Indels 146; Gaps 25;
 QY 30 CHCLEAPKCAPGVGLVRDGGCGCKVC---AKQLNEDCSKT-----QPCDHTKGLKC 79
 Db 558 CHM-LDLEVCVCSGLELYASLCAAGVCLPWRSHNTNTCTCTCPENQVYQPCGPNPHYCY 616
 QY 80 FGASSTALKGICRA--QSEGRPCENSHIYQ-----NGESFQP----- 115
 Db 617 RNDDISLSLAIQKAGPKSEGCFCDDMTLFSNDSICVPSQCQWCLGPHGPFVEFGHTISI 676
 QY 116 NCKHQCCTIDGAVGC-IPLCQBELSLNLCENPRLVKVS-----GCCBHWCEDESIK 169
 Db 677 NCQ-DCICKEGTLCQEKLCFQPT-----CPEGFVPSVIALEAGCCQFSC-----VC 725
 QY 170 DSLDDQDDLLGLDASEVELTRNNEILAIKKG--SSLKRLPVFGTEPRVLNPN-----LHAHG 224
 Db 726 NSSHCPPLHCPSESSLIVTEBGTCCPSQNSCKGKGVNGT-----LYQPGDVVSSSLC 781
 QY 225 QKCTVQRTS-----WSQCSKSGTGISTRV-----NDNPEC 256
 Db 782 ERCLCEVSSNAFSDVFNVCETELCNQCPKGFYQTTPGHCCGQCVPKTKPPKNSNST 841
 QY 257 RLVRK-----TRICE-----VR-----PCGQP-----VYSLK 279
 Db 842 SLYKPGFEPWEPNCPVTHKCEKFDVLTVTVTKIECPKINCPQDMAQLRGGDCYDCLV 901
 QY 280 KGKCKSKTKSPBPVRYTACSSVKYRKYK-CGCVDG-----RCCTP 323
 Db 902 PQCKTQVHQRO-QIIR--QONCSSEGPVSLSYCGNCGDSTSMYSLEANTVEHTCECCQE 959
 QY 324 LQTRTVKMRFRCEDEGMFSKNVMMIQSKCK-NVNCNPHNEAS 364
 Db 959 LQTSQRSVTLHCDGSSRTSFTYQVEKCGCLGQRCHAPGDT 1000

RESULT 9
 T03099
 Mucin, submaxillary - pig
 N;Alternate names: apomucin
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Jun-2000
 C;Accession: T03099; A40009; A28528; B29789
 R;Eckhardt, A.E.; Timpote, C.S.; Deluca, A.W.; Hill, R.L.
 J. Biol. Chem. 272, 33204-33210, 1997
 A;Title: The complete cDNA sequence and structural polymorphism of the polypeptide chain
 A;Reference number: Z14839; MUID:98070526; PMID:9407109
 A;Accession: T03099
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-13288 <ECK>
 A;Cross-references: EMBL:AF005273; NID:G2581863; PIDN:AA62527.1; PID:G2581864
 R;Eckhardt, A.E.; Timpote, C.S.; Abernethy, J.L.; Zhao, Y.; Hill, R.L.
 J. Biol. Chem. 266, 9678-9686, 1991
 A;Title: Porcine submaxillary mucin contains a cysteine-rich, carboxyl-terminal domain i
 A;Reference number: A40009; MUID:91236743; PMID:2033060
 A;Accession: A40009
 A;Molecule type: mRNA
 A;Residues: 12139-12167, 'T', 12169-13288 <EC3>
 A;Cross-references: GB:M61883; NID:G454837; PIDN:AAA30998.1; PID:G164374
 R;Timpote, C.S.; Eckhardt, A.E.; Abernethy, J.L.; Hill, R.L.
 J. Biol. Chem. 263, 1081-1088, 1988
 A;Title: Porcine submaxillary gland apomucin contains tandemly repeated, identical sequ
 A;Reference number: A28528; MUID:88087170; PMID:2826455
 A;Accession: A28528
 A;Molecule type: mRNA
 A;Residues: 12139-12167, 'T', 12169-12641 <TIM>
 A;Cross-references: GB:M21174; GB:J03512; NID:G164321; PIDN:AAA30990.1; PID:G552360
 A;Experimental source: submaxillary gland
 R;Eckhardt, A.E.; Timpote, C.S.; Abernethy, J.L.; Tounmadje, A.; Johnson Jr., W.C.; Hill,

J. Biol. Chem. 262, 11339-11344, 1987
A:Title: Structural properties of porcine submaxillary gland apomucin.
A:Reference number: A92606; MUID:87280230; PMID:3611111
A:Accession: B29789
A:Molecule type: protein
A:Residues: 1572-1607 <EC>
C:Superfamily: Pig submaxillary mucin
C:Keywords: tandem repeat

Query Match
Best Local Similarity 20.1%; Pred. No. 0.013;
Matches 93; Conservative 46; Mismatches 148; Indels 175; Gaps 22;

QY 26 CPAACHCPLKPCAPGVGLVDRGCGCKKCAKQALNEDCSKTPQCDHTKGLCNFGASST 85
DB 862 CKRGVCYP-----VGMWNSKGN-----VFPEDC-----886
QY 86 ALKGICRAQSEGRPCPEYNSRIYONGESFPNCKHQTCTIDGAVGCIP-----132
DB 887 -----PCSFGRGYDQGSVTSVGC-NKCTCIKGSWNCTQNECQTTCHYEGE 932
QY 133 -----LCPQELSLPNLGCNP--RLVKVSGQCEWVCDSDIKSLDDOD 176
DB 933 HIRTFDGTYSFDGLCQVSLFDYCGSENGTFRILTESVPCEDGLTCSRKIIIVAFQDN 992
QY 177 DLL-----GLDASEVELTRNNELI-AIGKGSLSKRLPVF-----GTEPRVLENPL 220
DB 993 VVLHDGKVTAVKTESKECELENSYSVHTVGLYLILKFLSGITIIIMDKNTRISVILDP- 1051
QY 221 HAHQKQ-----IVTTSWSQSKSCGIGTSTRVTDNPE 255
DB 1052 RWNGKVGCLGNNGDLKDDFTTRYSSVAVGELEFGNSWKTSQBCSDTVTQSPDCSNPY 1111
QY 256 CRLVKETRIICEV-----RPGQPVYSS-----LKKGKKSKTKKSPPEVRFTYAG-CS 302
DB 1112 CK-AWAVRKEIIRSTFRDCHNKVDPSAYDYACIEEACADMECK-----YLGECT 1162
QY 303 SVKKY-----RPKYCGSCVD-----GRCQ-----TPQTRVTKMRFCEDGEM 340
DB 1163 AVAMAEACSAVGCVTKWRKPLCLVPCYDYNAPGECWSRYEPCGTVTAK---TKORVI 1219
QY 341 FSKVNMWIOSC--KNYNCPHNEASERLYSL-----FNDI 374
DB 1220 GQKFSALLGCGVAKCPDSNPYLDENMKCVLSSESCFYNDI 1261

RESULT 10
T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089; PMID:9693030
A:Accession: T13954
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NA>
A:Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BA332462.1; P:G3449294
A:Experimental source: strain Sprague-Dawley; brain
C:Genetics:
A:Gene: MEGF6

Query Match
Best Local Similarity 20.4%; Score 156.5; DB 2; Length 1574;
Matches 94; Conservative 40; Mismatches 142; Indels 185; Gaps 24;

QY 26 CPAACHCPLKPCAPGVGLVDRGCGCKKCAKQALNEDCSKTPQCDHTKGLCNFGASST 85
DB 568 CSSPCTCQNGGTCDPVLGACRCPPGVSAHCEGCGPKFYGKHCKKCHCANRGRCHRLY 627

QY 51 GCKVKAKQL-----NEDSKTPQCDHTKGLCNFGASSTALKGICR 92
DB 628 GAC-LCDPGLYGRFCHLACPPWAFGPGCSBDCLEQ--SHTRSCNPKDSCS-----CK 678
QY 93 AQSEGRPCPEYNSRIYONGESFPNCKHQTCTIDGAVGCIP-----CPO 136
DB 679 AGFGERCQABC---ESG-FFGPGCRHRTCTQPG-VACDPVSGECRTQCPGPGVQGEDCG 733
QY 137 ELSLPNLG-----CPNRLVKVSGQC-CREWVCDSDIKSLDDQDDLLGLDASEVE 187
DB 734 ECPVGTGFGVNCSGSCGAPCHRVGTGECLECPKGTGDC---GADCPGGRWGLGQEI- 789
QY 188 LTRNNELIAGKSSLSKRLPVFVGTETPRVLPNPLHAHQKCIQVTT-----WSQCSK 239
DB 790 -----CPACEHGASCNPETGTCLECLPFGVGRQCQD 819
QY 240 SC-----GTGISR--VTNDN-----PECLVKETRIC-----VRPC-- 270
DB 820 TCSAGWYGTGQIRCAANDGHCDPTTGRCSAPGWTGLSCQRACDSHGWGPDCHPCNC 879
QY 271 -----GQPVVSSLKKGKSKTKKSPPEVRFTYAGSSVKKYRKYCGSC--VD 317
DB 880 SAGHGNDVAVSLGCLCEAGYEGPRCEQS-----CRQYVGFSCQKCRCEHGAACHVHS 933
QY 318 GRCTPLQTRTVKRFRCEDGEMFSKNVMMIOSCKNYNCP 358
DB 934 GACTCPAGWEGSPCEHACPAFPF---GLDCDSACNCSAGAP 971

RESULT 11
T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20336
A:Accession: T27283
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CESP:Y64G10A.f
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1;

Query Match
Best Local Similarity 23.2%; Score 156.5; DB 2; Length 1620;
Matches 92; Conservative 27; Mismatches 134; Indels 143; Gaps 22;

QY 26 CPAACHCPLKPCAPGVGLVDRGCGCKKCAKQALNEDCSK 66
DB 1052 CKGICSCQNGATCDSTVSGCECRPGWRKKCDRCPDG--RFEKGNAICDCTTTNDTSM 1109
QY 67 TQP-----CDHTKGLCNFGASSTALKGICRAQSEGRPCPEYNSRIYONGESFPNCKHQT 122
DB 1110 YNPFVACRCHVTG-ECR-----CPAGWTGDCQTSCTPLGRHGE-----GCRHSQ 1153
QY 123 CIDGA-----VGCIPLCQBELSLFNLGCPNRLVKVSGQCEEWVCD 165
DB 1154 CSNGASCDRVTFGDCPSGFMGNKCNCESECEPGLWSN--CMKHCLMHGECNKE-----1206
QY 166 DSIKSLDDQDDLLGLDASEVELTRNNELIAGKSSLSKRLPVFVGTETPRVLPNPLHAHQ 225
DB 1207 -----NGDCECIDGWTGSPCLPFGQFGNCAQRN--CKNGA 1241
QY 226 KCIQVQ-----TSWS--QCSKSGTGTSTRVTDNPECLVKETRICRVRPCGQPVYSS 277
DB 1242 SCDRKTGRCECLPGWSEHCEKSCVSG-----HYGAKC---EETCECEGALCDPISGH 1292
QY 278 LK-----KGGKSKTKKSPPEVRFTYAGSSVKKYRKYCGSC--VDGR-----320

Db 1293 CSCQPGWGRKKNR-----PCLKGYFGRHCSQCEKANSKCDHISGRCCQCPKGYAGHS 1346
 QY 321 CTPLQTRTVKMRPRCDGEMFKNVNMIOCKKNYN 356
 Db 1347 CTEL-----CPDG-TFGSCS--QKCDGCGEN 1369

RESULT 12
 T42626
 secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)
 N;Alternate names: neurogenic extracellular slit protein
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2002
 C;Accession: T42626
 R;Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.
 Mech. Dev. 79, 57-72, 1998
 A;Title: Distinct but overlapping expression patterns of two vertebrate slit homologs in
 A;Reference number: Z22177; MUID:99279238; PMID:10349621
 A;Accession: T42626
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1025 <HOL>
 A;Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1
 C;Genetics:
 A;Gene: Slit2
 C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein 2

Query Match 7.4%; Score 156; DB 2; Length 1025;
 Best Local Similarity 18.5%; Pred. No. 0.0018;
 Matches 101; Conservative 47; Mismatches 150; Indels 248; Gaps 23;

QY 24 STC-----PAACHCPLAPKAPGVLVRDGGCCCKVCAKQNLNEDCSKTQPCDH----- 72
 Db 507 STCVGDGINNYTCLCPPEYTG-----ELCEBKLDFAQDLNPPCHDSKCIL 551
 QY 73 -TKGLECNFGASGALKIGICRAQSEGRPCYNSRIYQNGESFQPNCKHOCTCIDGAVGCI 131
 Db 552 TPXGPKD-----CTPYIGEHCDIDFDDCQDNK-----CKNGAHTCDVANGYT 595
 QY 132 PLCPQ-----ELSLP-----NLGCPNPR--LVKVS-----GQCCREW 161
 Db 596 CVCPEGYSGLFCEFPFVPLPRTSPCDNFDCQNGAQCIIRINEPICQCLPGLVGEKCEKL 655
 QY 162 VCDEDSIKDS-----LDDQDDLGL-----DASEVELTRNNELTAT 197
 Db 656 VSNFVNKESYLQIPSAKVRPQNTILOATDEDSILLYKGDKHIAVELYRGRVASY 715
 QY 198 GKGS-----SLKRLPVFTEPRVLN-----PL 220
 Db 716 DTGSHPASAIYSVETINDGNFHIVELLTLDSSLSLSDVGGSPKVIITLSKQSTLNFDSPL 775
 QY 221 HA-----HG-----OKCIVQ 230
 Db 776 YVGMFGKNVASLRQAPGQNGTSFHGCTINLYINSELQDFRKMFMQGITLPGCFCHK 835
 QY 231 TTSWSQCSKSCGTGISTRV-----TND----- 252
 Db 836 VCAHGMQPSQSGFCECEBGMWPLCDQRTNDPCLGKNCVHGTCGLPNAFYSCKCLE 895
 QY 253 -----NPECLRVK--ETRLCEVRPCGPV--YSLKKGKCKSK--TKKSPFV 294
 Db 896 GHGVLDEEDLFPN--CQMIKCKHKKRLSGVGPYCECNSTGSDCDRISCRGERI 954
 QY 295 R-----FTYAGCSVKKYRKYK--GSCVDGRCTPLQTRTVKMRPRCEDGEMFKNVNM 347
 Db 955 RDYVQKQGYAACQTTKKVSRLECRGCGAGGCGGCGPLRSKRYKYSPECTDGSFVDEVEK 1014
 QY 348 IQSKCK 353
 Db 1015 VVRGCG 1020

RESULT 13
 A57534
 mucin SAC (clone L31) - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000
 C;Accession: A57534
 R;Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum, J. Biol. Chem. 270, 13665-13673, 1995
 A;Title: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cel
 A;Reference number: A57534; MUID:95293957; PMID:7775418
 A;Accession: A57534
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1042 <LES>
 A;Cross-references: GB:248314; NID:g1052607; PIDN:CAA88307.1; PID:g1052608
 C;Genetics:
 A;Gene: GDB:MUC5AC
 A;Cross-references: GDB:454136; OMIM:158373
 A;Map position: lip15.5-rip15.5
 C;Superfamily: von Willebrand factor type C repeat homology
 F:678-746/Domain: von Willebrand factor type C repeat homology <VMC>

Query Match 7.3%; Score 152.5; DB 2; Length 1042;
 Best Local Similarity 18.9%; Pred. No. 0.0032;
 Matches 105; Conservative 48; Mismatches 159; Indels 245; Gaps 27;

QY 4 STFTLAVAVTLHLTLRLALSTCPAACHCPLAPKAPGVLVRDGC-----GC 52
 Db 411 TVGSGTTVGPTTVGTTGPTTPAPC---LPSPICHLILSKVFEPCHTIVIPLLFYEGC 467
 QY 53 -----CKVCAK-----OLNEDCSKT-----QPCDHTKGL 76
 Db 468 VFDECHMTDLDVVCSLSLEYALCALASHDIDWRGTGHCPTCPADKVVYQCGSPNS 527
 QY 77 ECFNGASTALKGICRAQ--SEGRPCYNSRIYON-----GE 111
 Db 528 YC-YGNDASALGALPEAGPITEGCFPEGMTLFTSAQVCVPTGPRCLGPHGEPVKVGH 586
 QY 112 SFQPNCKHQCTC--IDGAVGCI--LCPQELSLNLGCPNRLVKV-----SGOCCEWVC 163
 Db 587 TVGMDCC--ECTCEAAWTTLTCRPLCP-----LPP--ACPLGFPVFPAPAGQCCPQYSC 640
 QY 164 DEDSIKSLDDQDDLGLDASEVELTRNNELIAGKSSILKRLV--FGTEPRVLFNPLHA 222
 Db 641 -----ACNTRSCPAPVGCPEGARAIPTYQ 664
 QY 223 HGQKCIYVTTWSQCSKSCGTGISTRVNDNPNCLVK-----ETRICE 266
 Db 665 EGACCPVQNCSTVTCISIN--GTLVQPGAVVSSSLCETCCBELPGGPPSDAFVWSCETQICN 723
 QY 267 V-----RPGQGPV-----YSSLKKG-----KCKSK-- 286
 Db 724 THCPVGFYEQSQCGCTCVQVACVTNTSKSPAHLFPYCGTWSDAGNHCVTHQCEKHQD 783
 QY 287 -----TKKSPFV-----RF-----TYAGCS 302
 Db 784 GLVVVTTTKACPLSLCSLDEARMSKDGCCRCPLPPPPYQNSTCAVYHRSLLIQOQCS 843
 QY 303 SVKKYRKYK--GSCVDG-----RCCTPLQTRTVKMRPRCEDGEMFKNVNM 346
 Db 844 SSEPVRLAYCRNGCGDSSSMYSLEGNTVHEHRCQCCQELRTSLNVLHCTDGSRAFSYT 903
 QY 347 MIQSKCK--NTNCPHPNE 362
 Db 904 EVEECGMGRRCFAPGD 920

RESULT 14
 A53767
 mucin MUC5B, tracheobronchial - human
 C;Species: Homo sapiens (man)
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000
 C;Accession: A53767

Search completed: April 22, 2004, 17:59:21
Job time : 19.4566 secs

Search completed: April 22, 2004, 17:59:21
Job time : 19.4566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:52:34 ; Search time 10.4724 Seconds
(without alignments)
1884.444 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103
Sequence: 1 MSSSTFRLAVAVTLHLTR.....PNEASFRLYSLFNDIHKFRD 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2103	100.0	379	1	CYR6_MOUSE
2	2079	98.9	379	1	CYR6_MOUSE
3	1938	92.2	381	1	CYR6_MOUSE
4	1646	78.3	375	1	CS10_CHICK
5	960	45.6	347	1	CTGF_MOUSE
6	957.5	45.5	348	1	CTGF_MOUSE
7	950	45.2	349	1	CTGF_MOUSE
8	937	44.6	349	1	CTGF_MOUSE
9	919	43.7	349	1	CTGF_MOUSE
10	861	40.9	351	1	NOV_CHICK
11	856	40.7	353	1	NOV_CHICK
12	839.5	39.9	357	1	NOV_CHICK
13	829.5	39.4	351	1	NOV_CHICK
14	827	39.3	354	1	NOV_CHICK
15	824.5	39.2	343	1	NOV_CHICK
16	773	36.8	367	1	WSP1_MOUSE
17	758.5	36.1	367	1	WSP1_MOUSE
18	757	36.0	367	1	WSP1_MOUSE
19	622	29.6	354	1	WSP2_MOUSE
20	526.5	25.0	251	1	WSP2_MOUSE
21	524	24.9	250	1	WSP2_MOUSE
22	492	23.4	250	1	WSP2_MOUSE
23	166	7.9	1529	1	SLT2_HUMAN
24	158	7.5	447	1	NRL1_MOUSE
25	154	7.3	456	1	NRL1_MOUSE
26	154	7.3	3110	1	LNA2_HUMAN
27	152.5	7.3	1233	1	MUSA_HUMAN
28	151.5	7.2	1700	1	BAR3_CHITE
29	148	7.0	450	1	NRL1_MOUSE
30	146	6.9	837	1	MUCL1_MOUSE
31	145.5	6.9	2813	1	WVF_CANFA
32	144.5	6.9	5703	1	MUSB_HUMAN
33	144	6.8	4655	1	LRP2_HUMAN

RESULT 1

ID	CYR6_MOUSE	STANDARD	PRT	379 AA
AC	P18406;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)			
DE	(Insulin-like growth factor-binding protein 10) (3CH61).			
GN	CYR61 OR IGFBP10 OR CCN1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Fibroblast;			
RX	MEDLINE=90287146; PubMed=2355916;			
RA	O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;			
RT	"Expression of cyr61, a growth factor-inducible immediate-early			
RT	gene";			
RL	Mol. Cell. Biol. 10:3569-3577(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A/J; TISSUE=Embryonic fibroblast;			
RX	MEDLINE=91288203; PubMed=2062642;			
RA	Latinkic B.V., O'Brien T.P., Lau L.F.;			
RT	"Promoter function and structure of the growth factor-inducible			
RT	immediate early gene cyr61";			
RL	Nucleic Acids Res. 19:3261-3267(1991).			
RN	[3]			
RP	HEPARIN-BINDING DOMAIN.			
RX	MEDLINE=20387398; PubMed=10821835;			
RA	Chen N., Chen C.C., Lau L.F.;			
RT	"Adhesion of human skin fibroblasts to Cyr61 is mediated through			
RT	integrin alpha 6beta 1 and cell surface heparan sulfate			
RT	proteoglycans";			
RL	J. Biol. Chem. 275:24953-24961(2000).			
CC	-/- FUNCTION: Promotes cell proliferation, chemotaxis, angiogenesis			
CC	and cell adhesion. Appears to play a role in wound healing by up-			
CC	regulating, in skin fibroblasts, the expression of a number of			
CC	genes involved in angiogenesis, inflammation and matrix remodeling			
CC	including VEGF-A, VEGF-C, MMP1, MMP3, TIMP1, uPA, PAI-1 and			
CC	integrins alpha-3 and alpha-5 (By similarity). Cyr61-mediated gene			
CC	regulation is dependent on heparin-binding (By similarity). Down-			
CC	regulates the expression of alpha-1 and alpha-2 subunits of			
CC	collagen type-1 (By similarity). Promotes cell adhesion and			
CC	adhesive signaling through integrin alpha-6/beta-1, cell migration			
CC	through integrin alpha-1/beta-5 and cell proliferation through			
CC	integrin alpha-v/beta-3 (By similarity).			
CC	-/- SUBUNIT: Interaction with integrins is heparin- and cell-type-			
CC	dependent and promotes cell adhesion (By similarity).			
CC	-/- SUBCELLULAR LOCATION: Secreted.			
CC	-/- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,			
CC	IN OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST			
CC	IN LUNG.			

ALIGNMENTS

	143	6.8	2482	1	VWF_PIG
34	141.5	6.7	1172	1	LMB3_HUMAN
35	141	6.7	1587	1	LMG3_HUMAN
36	141	6.7	5376	1	ZAN_MOUSE
37	141	6.7	5376	1	ZAN_MOUSE
38	140.5	6.7	2282	1	ZAN_RABBIT
39	140	6.7	3106	1	LMA2_MOUSE
40	139.5	6.6	1178	1	TSP2_CHICK
41	139	6.6	473	1	FP2_MYGA
42	138	6.6	1964	1	NTC4_MOUSE
43	136.5	6.5	5179	1	MUC2_HUMAN
44	136	6.5	480	1	HRA1_MOUSE
45	135.5	6.4	305	1	IBP2_MOUSE

Q28833 sus scrofa
Q13751 homo sapien
Q9Y6N6 homo sapien
O88799 mus musculus
P57399 oryctolagus
Q60675 mus musculus
P35440 gallus gall
Q25440 mytilus gall
P31695 mus musculus
Q02817 homo sapien
Q9R118 mus musculus
P47877 mus musculus


```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSPI; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSPI; 1.
DR PROSITE; PS01208; WVC; 1.
DR PROSITE; PS0184; WVC_2; 1.
KW Chemotaxis; Cell adhesion; Growth factor binding; Heparin-binding;
FT SIGNAL 1 24
FT CHAIN 25 379
FT DOMAIN 26 97
FT DOMAIN 98 164
FT DOMAIN 226 271
FT DOMAIN 277 313
FT DOMAIN 284 358
FT DISULFID 284 321
FT DISULFID 301 335
FT DISULFID 312 351
FT DISULFID 315 353
FT DISULFID 320 357
SQ SEQUENCE 379 AA; 41687 MW; 62BF03BA4C5AFDE9 CRC64;

Query Match 98.9%; Score 2079; DB 1; Length 379;
Best Local Similarity 98.7%; Pred. No. 1.7e-154;
Matches 374; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSFTFRLAVAVTLHLTRALSTCPAACHCPLEAPKAPGVLVDRDGGCKVKCAKOL 60
DB 1 MSSFTIKTAVAVTLHLTRALSTCPAACHCPLEAPKAPGVLVDRDGGCKVKCAKOL 60
QY 61 NEDCSKTQPCDHTTGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFPQCKHQ 120
DB 61 NEDCSKTQPCDHTTGLGECNFGASSTALKGICRAQSEGRPCPEYNSRIYQNGESFPQCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNCPPRLVKVSGQCEWDCEDSIKSLDDQDQLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNCPPRLVKVSGQCEWDCEDSIKSLDDQDQLLG 180
QY 181 LDASEVELTRNNELIAIGKSSLRPLVFGTEPRVLFNPLHAGQKCIQVTTWSQCSKS 240
DB 181 FDASEVELTRNNELIAIGKSSLRPLVFGTEPRVLYNPLHAGQKCIQVTTWSQCSKS 240
QY 241 CGTGISTRTVNDNEECRLVKETRICVVRPCGQPVYSSLLKGGKSKTKKSPVRYTYAG 300
DB 241 CGTGISTRTVNDNEECRLVKETRICVVRPCGQPVYSSLLKGGKSKTKKSPVRYTYAG 300
QY 301 CSSVKYRKYKCGSDVGRCTPLQTRIVKVRFCEDGEMFSKNVMVMIQCKNYNCNCPH 360
DB 301 CSSVKYRKYKCGSDVGRCTPLQTRIVKVRFCEDGEMFSKNVMVMIQCKNYNCNCPH 360
QY 361 NEASFRLYSLENDIHKFRD 379
DB 361 NEASFRLYSLENDIHKFRD 379

RESULT 3
CYR6 HUMAN STANDARD; PRT; 381 AA.
AC O0622; O14934; O43775; Q9BZL7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10) (GIGI protein).
```

```
GN CYR61 OR IGFBP10 OR CCN1 OR GIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=97280750; PubMed=9135077;
RA Jay P., Berge-LeFranc J.L., Marsoiller C., Mejean C., Taviaux S.,
RA Berta P.;
RT "The human growth factor-inducible immediate early gene, CYR61, maps
RT to chromosome 1p";
RL Oncogene 14:1753-1757(1997).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=98197344; PubMed=9536281;
RA Martineze C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
RT cells from the nervous system.";
RL Mol. Pathol. 50:310-316(1997).
RN [4]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kolesnikova T.V., Lau L.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA Bi A.B., Yu L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RA Schuetz N., Lechner A., Groll C., Koehle J., Jakob F.;
RT "Regulation of hCYR61 by vitamin D, serum and cytokines in fetal human
RT osteoblasts.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
RT "Organization and expression of the CYR61 gene in normal human
RT fibroblasts.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RN SEQUENCE FROM N.A.
RC TISSUE=Lung, Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whitling R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RN FUNCTION IN WOUND HEALING.
RX MEDLINE=21601638; PubMed=11584015;
RA Chen C.-C., Mo F.-E., Lau L.F.;
```

RT "The angiogenic factor Cyr61 activates a genetic program for wound healing in human skin fibroblasts.";
 RL J. Biol. Chem. 276:47329-47337(2001).
 RN [10]
 RP INTERACTION WITH INTEGRIN ALPHA-V/BETA-3.
 RX MEDLINE=98112865; PubMed=9446626;
 RA Kireeva M.L., Lam S.C., Lau L.F.;
 RT Adhesion of human umbilical vein endothelial cells to the immediate-early gene product Cyr61 is mediated through integrin alphavbeta3.";
 RL J. Biol. Chem. 273:3090-3096(1998).
 RN [11]
 RP INTERACTION WITH INTEGRIN ALPHA-IIB/BETA-3.
 RX MEDLINE=99377072; PubMed=10446209;
 RA Jadesayannata A., Chen C.-C., Kireeva M.L., Lau L.F., Lam S.C.;
 RT Activation-dependent adhesion of human platelets to Cyr61 and Fsp1z/mouse connective tissue growth factor is mediated through integrin alpha(iiib)beta(3).";
 RL J. Biol. Chem. 274:24321-24327(1999).
 RN [12]
 RP INTERACTION WITH INTEGRIN ALPHA-6/BETA-1.
 RX MEDLINE=20387398; PubMed=10821835;
 RA Chen N., Chen C.-C., Lau L.F.;
 RT Adhesion of human skin fibroblasts to Cyr61 is mediated through integrin alphabeta1 and cell surface heparan sulfate proteoglycans.";
 RL J. Biol. Chem. 275:24953-24961(2000).
 RN [13]
 RP INTERACTION WITH INTEGRIN ALPHA-V/BETA-6.
 RX MEDLINE=21293087; PubMed=11287419;
 RA Grzeszkiewicz T.M., Kirschling D.J., Chen N., Lau L.F.;
 RT "Cyr61 stimulates human skin fibroblast migration through Integrin alpha vbeta 5 and enhances mitogenesis through integrin alpha vbeta 3, independent of its carboxyl-terminal domain.";
 RL J. Biol. Chem. 276:21943-21950(2001).
 CC -1- FUNCTION: Promotes cell proliferation, chemotaxis, angiogenesis and cell adhesion. Appears to play a role in wound healing by up-regulating in skin fibroblasts, the expression of a number of genes involved in angiogenesis, inflammation and matrix remodeling including VEGF-A, VEGF-C, MMP1, MMP3, TIMP1, uPA, PAI-1 and integrins alpha-3 and alpha-5. Cyr61-mediated gene regulation is dependent on heparin-binding. Down-regulates the expression of alpha-1 and alpha-2 subunits of collagen type-1. Promotes cell adhesion and adhesive signaling through integrin alpha-6/beta-1, cell migration through integrin alpha-v/beta-5 and cell proliferation through integrin alpha-v/beta-3.
 CC -1- SUBUNIT: Interaction with Integrins is heparin- and cell-type-dependent and promotes cell adhesion. In skin fibroblasts it binds ITGA6/ITGB1, in endothelial cells, binds ITGA5/ITGB3 and in platelets, ITGA2B/ITGB3. Binds, in vitro, ITGA5/ITGB5.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the CCN family.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL; Y12084; CAA72802.1; -;
 DR EMBL; U62015; AAB58319.1; -;
 DR EMBL; Y11307; CAA72167.1; -;
 DR EMBL; AF003594; AAB61240.1; -;
 DR EMBL; AF031385; AAB84227.1; -;
 DR EMBL; Z98053; CAB10848.1; -;
 DR EMBL; AF307860; AAG59863.1; -;
 DR EMBL; BC001271; AAH01271.1; -;
 DR EMBL; BC009199; AAH09199.1; -;

DR EMBL; BC016952; AAH16952.1; -;
 DR Genew; HGNC:2654; CYR61.
 DR MIM; 602369; -;
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSPI.
 DR InterPro; IPR001007; WFC_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp.1; 1.
 DR Pfam; PF00093; wfc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSPI; 1.
 DR SMART; SM00214; WFC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00232; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSPI; 1.
 DR PROSITE; PS01208; WFC_1; 1.
 DR PROSITE; PS0184; WFC_2; 1.
 DR Chemotaxis; Cell adhesion; Growth factor binding; Heparin-binding; KW
 Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 381 CYR61 PROTEIN.
 FT DOMAIN 96 97 IGFBP.
 FT DOMAIN 228 273 TSP TYPE-1.
 FT DOMAIN 279 315 HEPARIN-BINDING (BY SIMILARITY).
 FT DOMAIN 286 360 CTCK.
 FT DISULFID 286 323 BY SIMILARITY.
 FT DISULFID 303 337 BY SIMILARITY.
 FT DISULFID 314 353 BY SIMILARITY.
 FT DISULFID 317 355 BY SIMILARITY.
 FT DISULFID 322 359 BY SIMILARITY.
 FT CONFLICT 165 165 E -> Q (IN REF. 3).
 FT CONFLICT 210 210 L -> I (IN REF. 5).
 FT CONFLICT 220 220 L -> R (IN REF. 5).
 FT CONFLICT 369 369 F -> L (IN REF. 7).
 SQ SEQUENCE 381 AA; 42026 MW; FC0BD39C078CA0B1 CRC64;
 Query Match 92.2%; Score 1938; DB 1; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.5e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 Qy 1 MSSSTRTFLAVVTLHLRLALSTCPAACHCPLEAPKCAPGVGLVDRDGCCKVCAKQL 60
 Db 1 MSSRIARALAVTLHLRLALSTCPAACHCPLEAPKCAPGVGLVDRDGCCKVCAKQL 60
 Qy 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAQSEGRCEYNSRIYONGESFPQCKHQ 120
 Db 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAQSEGRCEYNSRIYONGESFPQCKHQ 120
 Qy 121 CTGIDGAVGICPLCPBELSLPNLGNCNPNRLVKVSGCCCEWVCDSDSINDSLDDQDDL-- 178
 Db 121 CTGIDGAVGICPLCPBELSLPNLGNCNPNRLVKVSGCCCEWVCDSDSINDSLDDQDGLLG 180
 Qy 179 --LGLDASEVELTRNNELIAHGKSLKPLVPFGTEPRVLFNPLHAGKQKICVQTTSWSQ 236
 Db 181 KELGFDASEVELTRNNELIAVGKSSLKPLVPFGMEPRILYNPL--QGKQKICVQTTSWSQ 238
 Qy 237 CSKSCGTGISTRTVNDNPECLVKETRICEVPRCPGQPVYSSLLKKGKCKSKTKKSPFVR 296
 Db 239 CSKTGCTGISTRTVNDNPECLVKETRICEVPRCPGQPVYSSLLKKGKCKSKTKKSPFVR 298
 Qy 297 TYAGCSVKKYRKYCGSCVDGRCCTPLQTRTVVMFRFRDGMFNSNVMMIOSCKNYN 356
 Db 299 TYAGCLSVKKYRKYCGSCVDGRCCTPQLTRTVVMFRFRDGMFNSNVMMIOSCKNYN 358
 Qy 357 CPHENASFRLYSLFNDIHKFRD 379

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	DB 3	DB 4	DB 5	DB 6	DB 7	DB 8	DB 9	DB 10	DB 11	DB 12	DB 13	DB 14	DB 15	DB 16	DB 17	DB 18	DB 19	DB 20	DB 21	DB 22	DB 23	DB 24	DB 25	DB 26	DB 27	DB 28	DB 29	DB 30	DB 31	DB 32	DB 33	DB 34	DB 35	DB 36	DB 37	DB 38	DB 39	DB 40	DB 41	DB 42	DB 43	DB 44	DB 45	DB 46	DB 47	DB 48	DB 49	DB 50	DB 51	DB 52	DB 53	DB 54	DB 55	DB 56	DB 57	DB 58	DB 59	DB 60	DB 61	DB 62	DB 63	DB 64	DB 65	DB 66	DB 67	DB 68	DB 69	DB 70	DB 71	DB 72	DB 73	DB 74	DB 75	DB 76	DB 77	DB 78	DB 79	DB 80	DB 81	DB 82	DB 83	DB 84	DB 85	DB 86	DB 87	DB 88	DB 89	DB 90	DB 91	DB 92	DB 93	DB 94	DB 95	DB 96	DB 97	DB 98	DB 99	DB 100	DB 101	DB 102	DB 103	DB 104	DB 105	DB 106	DB 107	DB 108	DB 109	DB 110	DB 111	DB 112	DB 113	DB 114	DB 115	DB 116	DB 117	DB 118	DB 119	DB 120	DB 121	DB 122	DB 123	DB 124	DB 125	DB 126	DB 127	DB 128	DB 129	DB 130	DB 131	DB 132	DB 133	DB 134	DB 135	DB 136	DB 137	DB 138	DB 139	DB 140	DB 141	DB 142	DB 143	DB 144	DB 145	DB 146	DB 147	DB 148	DB 149	DB 150	DB 151	DB 152	DB 153	DB 154	DB 155	DB 156	DB 157	DB 158	DB 159	DB 160	DB 161	DB 162	DB 163	DB 164	DB 165	DB 166	DB 167	DB 168	DB 169	DB 170	DB 171	DB 172	DB 173	DB 174	DB 175	DB 176	DB 177	DB 178	DB 179	DB 180	DB 181	DB 182	DB 183	DB 184	DB 185	DB 186	DB 187	DB 188	DB 189	DB 190	DB 191	DB 192	DB 193	DB 194	DB 195	DB 196	DB 197	DB 198	DB 199	DB 200	DB 201	DB 202	DB 203	DB 204	DB 205	DB 206	DB 207	DB 208	DB 209	DB 210	DB 211	DB 212	DB 213	DB 214	DB 215	DB 216	DB 217	DB 218	DB 219	DB 220	DB 221	DB 222	DB 223	DB 224	DB 225	DB 226	DB 227	DB 228	DB 229	DB 230	DB 231	DB 232	DB 233	DB 234	DB 235	DB 236	DB 237	DB 238	DB 239	DB 240	DB 241	DB 242	DB 243	DB 244	DB 245	DB 246	DB 247	DB 248	DB 249	DB 250	DB 251	DB 252	DB 253	DB 254	DB 255	DB 256	DB 257	DB 258	DB 259	DB 260	DB 261	DB 262	DB 263	DB 264	DB 265	DB 266	DB 267	DB 268	DB 269	DB 270	DB 271	DB 272	DB 273	DB 274	DB 275	DB 276	DB 277	DB 278	DB 279	DB 280	DB 281	DB 282	DB 283	DB 284	DB 285	DB 286	DB 287	DB 288	DB 289	DB 290	DB 291	DB 292	DB 293	DB 294	DB 295	DB 296	DB 297	DB 298	DB 299	DB 300	DB 301	DB 302	DB 303	DB 304	DB 305	DB 306	DB 307	DB 308	DB 309	DB 310	DB 311	DB 312	DB 313	DB 314	DB 315	DB 316	DB 317	DB 318	DB 319	DB 320	DB 321	DB 322	DB 323	DB 324	DB 325	DB 326	DB 327	DB 328	DB 329	DB 330	DB 331	DB 332	DB 333	DB 334	DB 335	DB 336	DB 337	DB 338	DB 339	DB 340	DB 341	DB 342	DB 343	DB 344	DB 345	DB 346	DB 347	DB 348	DB 349	DB 350	DB 351	DB 352	DB 353	DB 354	DB 355	DB 356	DB 357	DB 358	DB 359	DB 360	DB 361	DB 362	DB 363	DB 364	DB 365	DB 366	DB 367	DB 368	DB 369	DB 370	DB 371	DB 372	DB 373	DB 374	DB 375	DB 376	DB 377	DB 378	
-------------	-----------------------	-------	------	--------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--

CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF120275; AAD39132.1; -;
 CC EMBL; AB023068; BAA82125.1; -;
 CC InterPro; IPR006208; Cys knot.
 CC InterPro; IPR006207; Cys knot C.
 CC InterPro; IPR000857; InaI_gro_fac_pr.
 CC InterPro; IPR000884; TSPI.
 CC InterPro; IPR001007; WVF_C.
 CC Pfam; PF00007; Cys knot; 1.
 CC Pfam; PF00219; IGFBP; 1.
 CC Pfam; PF00090; tsp; 1; 1.
 CC Pfam; PF00093; vwc; 1.
 CC SMART; SM00041; CT; 1.
 CC SMART; SM00121; IS; 1.
 CC SMART; SM00209; TSPI; 1.
 CC SMART; SM00214; VWC; 1.
 CC PROSITE; PS01185; CTCK_1; 1.
 CC PROSITE; PS01225; CTCK_2; 1.
 CC PROSITE; PS00222; IGF BINDING; 1.
 CC PROSITE; PS00032; TSPI; 1.
 CC PROSITE; PS01208; VWC; 1; 1.
 CC PROSITE; PS0184; VWC_2; 1.
 CC Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
 KW Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 347
 FT DOMAIN 31 98
 FT DOMAIN 99 165
 FT DOMAIN 196 241
 FT DOMAIN 245 347
 FT DOMAIN 254 328
 FT DISULFID 254 291
 FT DISULFID 271 305
 FT DISULFID 282 321
 FT DISULFID 285 323
 FT DISULFID 290 327
 FT DISULFID 335 35
 FT CONFLICT 35 35 A -> R (IN REF. 2).
 FT CONFLICT 94 94 T -> P (IN REF. 2).
 SQ SEQUENCE 347 AA; 37756 MW; CFBELAI9766B7B16 CRC64;
 Query Match 45.6%; Score 960; DB 1; Length 347;
 Best Local Similarity 46.4%; Pred. No. 1.4e-67;
 Matches 176; Conservative 61; Mismatches 104; Indels 38; Gaps 7;
 QY 1 MSSSTFRTAVATLLHLRLAL-STCPAAACHCPLP-APKAPGVGLVRDGGCGCKVCVK 58
 DB 1 MLASVAGPVSLALVLLLTPTATGQCSAQCAAEAPRCAPAGVSLVLDGGCCRCVK 60
 QY 59 QLNECDKSTQPCDHTTGLNCFGASSTALKGICRAQSGRCPCVNSRIYONGESFPQCK 118
 DB 61 QLGLCTERDPCDPHGLGFCDFGSPANRKLGVCTAK-DGAPCVGSGVRSGESFQSSCK 119
 QY 119 HQCTIDGAGVCTPLCPQBELSLPNLCNPNRLVKVSGCCBWEVCDSDIKSLDDQDDL 178
 DB 120 YQCTCLDGAGVCPVLCMDVRLPSDPCFPFRVRLPGKCCBWEVCDP----- 167
 QY 179 LGLDASEVELTRNNELIATKGSSLRLL-PVFTPEPRVLFNPLHAHQKQCIQVTTWSQC 237
 DB 168 -----KDRTVGPALAAVRLDTPGDPFTM-----RANCLVQTTEWAC 207
 QY 238 SKSCGTGISTRVNDNPECLVKETRICVPRGQPVYSSLLKKGKCKTKKSPFPVFT 297
 DB 208 SKTCGMGISTRVNDNTFCRLKQSLCMLWVRPCEADLEENIKKGGKCIKTRPKIAPVKPE 267

QY 298 YAGCSSVKYRKYCGSCVDGRCCTPLQTRTKVMBPRCEDGEMFSKNVMIOSCKNYNC 357
 DB 268 LSGCTSVYTKYRAKFCGVCVDRGCCPHRTITLTFVFKCPDGEIMKKNMFIKTCACHYC 327
 QY 358 PHNEASFRILY--SLFNDI 374
 DB 328 PGDNDIFESLYRKMVGDM 346
 RESULT 6
 CTGF MOUSE
 ID CTGF MOUSE STANDARD; PRT; 348 AA.
 AC P29268; Q922U0;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Connective tissue growth factor precursor (FISP-12 protein)
 DE (Hypertrophic chondrocyte-specific protein 24).
 GN CTGF OR CCN2 OR FISP12 OR FISP-12 OR HCS24.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91363290; PubMed=1886698;
 RA Ryseck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
 RT "Structure, mapping, and expression of fisp-12, a growth factor-
 RT inducible gene encoding a secreted cysteine-rich protein.";
 RL Cell Growth Differ. 2:225-233(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91229699; PubMed=2029337;
 RA Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
 RT "Identification of a gene family regulated by transforming growth
 RT factor-beta.";
 RL DNA Cell Biol. 10:293-300(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=97327410; PubMed=9184077;
 RA Kireeva M.L., Latinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
 RA Abler A.S., Lau L.F.;
 RT "Cyr61 and fisp12 are both ECM-associated signaling molecules;
 RT activities, metabolism, and localization during development.";
 RL Exp. Cell Res. 233:63-77(1997).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=99182484; PubMed=10082563;
 RA Babic A.M., Chen C.C., Lau L.F.;
 RT "Fisp12/mouse connective tissue growth factor mediates endothelial

cell adhesion and migration through integrin alphavbeta3, promotes endothelial cell survival, and induces angiogenesis in vivo.";
 Mol. Cell. Biol. 19:2958-2966(1999).
 -!- FUNCTION: Major connective tissue mitogen secreted by vascular endothelial cells. Promotes proliferation and differentiation of chondrocytes (By similarity). Mediates heparin- and divalent cation-dependent cell adhesion in many cell types including fibroblasts, myofibroblasts, endothelial and epithelial cells (By similarity). Enhances fibroblast growth factor-induced DNA synthesis (By similarity).
 -!- SUBUNIT: Monomer (By similarity).
 -!- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a soluble form.
 -!- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 -!- INDUCTION: By growth factors.
 -!- SIMILARITY: Belongs to the CCN family.
 -!- SIMILARITY: Contains 1 IGFBP domain.
 -!- SIMILARITY: Contains 1 WFPC domain.
 -!- SIMILARITY: Contains 1 TSP type-1 domain.
 -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M70641; AAA37627.1; -
 EMBL; M70642; AAA37628.1; -
 EMBL; M80263; AAA73135.1; -
 EMBL; BC006783; AA06783.1; -
 PIR; A40578; A40578.
 MGD; MGI-95537; Ctgf.
 GO; GO:0005578; C:extracellular matrix; IDA.
 GO; GO:0008201; F:heparin binding; IDA.
 GO; GO:0005178; F:integrin binding; IDA.
 GO; GO:0001525; P:angiogenesis; IDA.
 GO; GO:0016477; P:cell migration; IDA.
 GO; GO:0007160; P:cell-matrix adhesion; IDA.
 GO; GO:0008543; P:FGF receptor signaling pathway; IDA.
 GO; GO:0007223; P:integrin-mediated signaling pathway; IDA.
 GO; GO:0001503; P:ossification; IMP.
 InterPro; IPR006208; Cys_knot.
 InterPro; IPR000867; Insl_gro_fac_pr.
 InterPro; IPR000884; TSP1_gro_fac.
 InterPro; IPR01007; VWF_C.
 Pfam; PF00007; Cys_knot; 1.
 Pfam; PF02119; IGFBP; 1.
 Pfam; PF00090; tsp_1; 1.
 Pfam; PF00093; vwc; 1.
 SMART; SM00041; CT; 1.
 SMART; SM00121; IB; 1.
 SMART; SM00209; TSP1; 1.
 SMART; SM00214; VWC; 1.
 PROSITE; PS01185; CTCK_1; 1.
 PROSITE; PS01225; CTCK_2; 1.
 PROSITE; PS02022; IGF_BINDING; 1.
 PROSITE; PS00092; TSP1; 1.
 PROSITE; PS01208; VWF_1; 1.
 PROSITE; PS01208; VWF_2; 1.
 PROSITE; PS01084; VWF_2; 1.
 Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding; Signal.
 KW SIGNAL.
 FT CHAIN 1 25
 FT CHAIN 26 348
 FT DOMAIN 32 99
 FT DOMAIN 100 166
 FT DOMAIN 197 242
 FT DOMAIN 246 348
 FT DOMAIN 255 329
 CTCK.

FT DISULFID 255 292 BY SIMILARITY.
 FT DISULFID 272 306 BY SIMILARITY.
 FT DISULFID 283 322 BY SIMILARITY.
 FT DISULFID 286 324 BY SIMILARITY.
 FT DISULFID 291 328 BY SIMILARITY.
 FT CONFLICT 161 161 E -> K (IN REF. 1).
 SQ SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;
 Query Match 45.5%; Score 957.5; DB 1; Length 348;
 Best Local Similarity 46.8%; Pred. No. 2.2e-67;
 Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;
 QY 1 MSSSTRTTAVAVTLLHL-TRLAL-STCPAAACHCPLE-APKCAPGVGLVDRGGCGCKVCA 57
 DB 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCCAAEAAPHCPAGVSLVLDGGCCRVCA 60
 QY 58 KQLNEDCSKTQPCDHTKGLNCFGASSTALKGICRAQSEGRPCSEYNSRIYQNGESFPQNC 117
 DB 61 KQLGELCTERDFCDPHKGLFCDFGSPANKIGVCTAK-DGAPCVFGSVYRSGESFQSSC 119
 QY 118 KQCTCIDGAVGICPLCPQELSLPNIUGCNPPLVNVKVGSCCEEWVCDENSIKSLDDQDD 177
 DB 120 KYQCTCLDGAAGCVPLCSMDVRLPSPDCFPFRVRLPGKCCSEWVCDPEP----- 168
 QY 178 LLGLDASEVELTRNNELIAIGKGSILKRL-PVFGTEPRVLFNPLHAHGOKCIVQTTWSQ 236
 DB 169 -----KDRTAGVAPALAAAYRLEDTFGDPTM-----RANCLVQTTWSA 207
 QY 237 CSKSGCTGTSTVNDNPECLRVKSTRICEVPGQGVYSSLKKGKCKSKTKKSPVRF 296
 DB 208 CSKTCGMGISTRTVNDNTFRCLEKQSLCWMVFCADLEENIKGKCKCIRTPIAKPVKF 267
 QY 297 TVAGSSVVKYRKYKCGSVDCRCCTPLQTRTVKMRFCEDCEMFSSKNVMTQSCKNYN 356
 DB 268 ELSGCTSVKTYRAXFCGVCTDRCCTPHRTTTLVPEFKCPDGEIMKKNMFKTCACHN 327
 QY 357 CHPNEASPRLY--SLFNDI 374
 DB 328 CFGDNDIFESLYYRKYMGDM 347
 RESULT 7
 CTGF_HUMAN
 ID CTGF_HUMAN STANDARD; PRT; 349 AA.
 AC P29279; Q960X2;
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 42, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Connective tissue growth factor precursor (Hypertrophic chondrocyte-specific protein 24).
 DE CTGF OR CCN2 OR HCS24.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Umbilical vein endothelial cells;
 RX MEDLINE=91373462; PubMed=1654338;
 RA Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
 RT "Connective tissue growth factor: a cysteine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CEF-10.";
 RL J. Cell Biol. 114:1295-1294(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Umbilical vein endothelial cells;
 RX MEDLINE=93187114; PubMed=1293144;
 RA Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
 RT "Connective tissue growth factor.";
 RL J. Dermatol. 19:642-643(1992).
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE=Aorta;
 RX MEDLINE=97207446; PubMed=9054739;
 RA Omar B.S., Warner A., Garnier J.M., Do D.D., Godoy N., Nauck M.,
 RA Marz W., Rupp J., Pech M., Luescher T.F.,
 RT "Human connective tissue growth factor is expressed in advanced
 RT atherosclerotic lesions";
 RL Circulation 95:831-839(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Copley V.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP FUNCTION.
 RC TISSUE=Chondrocytes;
 RX MEDLINE=20080284; PubMed=10614647;
 RA Nakanishi T., Nishida I., Shimo T., Kobayashi K., Kubo T.,
 RA Tamatani T., Tezuka K., Takigawa M.;
 RT "Effects of CTGF/Hcs24, a product of a hypertrophic chondrocyte-
 RT specific gene, on the proliferation and differentiation of
 RT chondrocytes in culture.";
 RL Endocrinology 141:264-273(2000).
 RN [6]
 RP HEPARIN-BINDING, AND CELL ADHESION.
 RX MEDLINE=22442376; PubMed=1253878;
 RA Ball D.K., Rachfal A.W., Kemper S.A., Brigstock D.R.;
 RT "The heparin-binding 10 kDa fragment of connective tissue growth
 RT factor (CTGF) containing module 4 alone stimulates cell adhesion.";
 RL J. Endocrinol. 176:R1-R7(2003).
 CC -!- FUNCTION: Major connective tissue microattractant secreted by
 CC vascular endothelial cells. Promotes proliferation and
 CC differentiation of chondrocytes. Mediates heparin- and divalent
 CC cation-dependent cell adhesion in many cell types including
 CC fibroblasts, myofibroblasts, endothelial and epithelial cells.
 CC Enhances fibroblast growth factor-induced DNA synthesis.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a
 CC soluble form (by similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Name=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Long;
 CC IsoId=P29279-1; Sequences=Displayed;
 CC Name=2; Synonyms=Short;
 CC IsoId=P29279-2; Sequences=VSP 002460;
 CC Note=No experimental confirmation available;
 CC -!- SIMILARITY: Belongs to the CCN family.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M29294; AA01279.1; -;
 DR EMBL; X78947; CA05544.1; -;
 DR EMBL; AL354866; CAC44023.1; -;
 DR FIR; A40551; A40551.
 DR Genew; HGNC:2500; CTGF.
 DR MIM; 121009; -;
 DR GO; GO:0005578; C:extracellular matrix; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:000520; F:insulin-like growth factor binding; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR GO; GO:0008544; P:epidermal differentiation; TAS.
 DR GO; GO:0009611; P:response to wounding; TAS.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Inei_gro_fac_pr.

DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; VWFC.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; VWFC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK 1; 1.
 DR PROSITE; PS01225; CTCK 2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00322; TSP1; 1.
 DR PROSITE; PS01208; VWFC; 1; 1.
 DR PROSITE; PS0184; VWFC2; 1.
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
 KW Signal; Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 349
 FT DOMAIN 33 100
 FT DOMAIN 101 167
 FT DOMAIN 198 243
 FT DOMAIN 247 349
 FT DOMAIN 256 293
 FT DISULFID 273 307
 FT DISULFID 284 323
 FT DISULFID 287 325
 FT DISULFID 292 329
 FT CARBOHYD 28 28
 FT CARBOHYD 225 225
 FT VARSPPLIC 172 198
 FT CONFLICT 83 83 D -> H (IN REF. 4).
 SQ SEQUENCE 349 AA; 38069 MW; 0ECF8470B357EA95 CRC64;
 Query Match 45.2%; Score 950; DB 1; Length 349;
 Best Local Similarity 45.7%; Pred. No. 8.3e-67;
 Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;
 QY 1 MSSSTFTTLLAVVTLHLRLALSTCPAA-----CHCPLP-APKCAPGVGLVRDGGCG 52
 DB 1 MTASMGPRVAVPVL-----LALCSRPVGVQNSGFCRCPDFAPRCPAGVSLVLDGGCG 56
 QY 53 CKYCAKQLNEDCSKTQPCDHTKGLCNFGASSTALKICRAOSEGRPCYNNRIYONGES 112
 DB 57 CRVCAKQLGELCTERDPCDPKGLFCDFGSPANRKIGVCTAK-DGAPCIFGGTVYRSGES 115
 QY 113 FQPNCKHCTCIDGAYGCIPLCPQELSLNLCNPRLVKVSQCCEWVCDSDSIKDSL 172
 DB 116 FQSCRYCQCTCLDGAAGCMPLCSMDVRLSPDCPPRRVRLPKCCCEWVCDDEPK----- 170
 QY 173 DDQDDLGLDASEVELTRNNEIAIGKSSKLRL-PVFGTEPRVLNPLHAHGKQCIQVOT 231
 DB 171 -DQ-----TVGPAALAAVRLEDTFGDPDTMI-----RANCLVQT 203
 QY 232 TSWSQCSKSGTGISTRTNDNPECLVKETICEVRCQGPVYSLLKKGKCKSKTKKSP 291
 DB 204 TEWSACSKCTGMISTRVNDNASCKLEKQSLCVMWRPCEADLEENIKKGKCKIRTPKLS 263
 QY 292 EPVRFYTAGSSVKYRPRKYCGSCVDGRCTCCTPLQTRTVKORFRCEDGENFSSKNVMMIQSC 351
 DB 264 KPIKFELSGCTSMKTYRAKFCGVCTDGRCTCCTPHRTTLLPVEFKCPDGEVKNMMFIKC 323
 QY 352 KNYNCPHNEASFRLY--SLFNDI 374
 DB 324 ACHNCPGNDIFESLYYRKMVGDM 348
 RESULT 8
 CTGF_BOVIN STANDARD; PRT; 349 AA.
 ID_CTGF_BOVIN

AC 018739; 09GL71;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Connective tissue growth factor precursor.
 GN CTGF OR CCN2.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aorta;
 RA Liliensiek B., Lin Z., Fotsis T., Schimanski M., Bierhaus A.,
 RA Kanitz M., Kauffmann G., Schweigert L., Ziegler R., Nawroth P.P.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Mathias M., Schwitters C., Hove M., Rupp S., Brundu N.E.;
 RT "Bovine connective tissue growth factor, organization of the
 RT chromosomal gene and demonstration of promoter activity";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Major connective tissue mitottractant secreted by
 CC vascular endothelial cells. Promotes proliferation and
 CC differentiation of chondrocytes (By similarity). Mediates heparin-
 CC and divalent cation-dependent cell adhesion in many cell types
 CC including fibroblasts, myofibroblasts, endothelial and epithelial
 CC cells (By similarity). Enhances fibroblast growth factor-induced
 CC DNA synthesis (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Found in the extracellular matrix and as a
 CC soluble form (By similarity).
 CC -!- SIMILARITY: Belongs to the CCN family.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 WFEC domain.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF000137; AAB66596.1; -;
 DR EMBL; AF309555; AAG30290.1; -;
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Inel_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00203; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS01208; VWF_1; 1.
 DR PROSITE; PS0184; VWF_2; 1.
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Heparin-binding;
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 26 POTENTIAL.
 FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.

FT	DOMAIN	33	100	IGFBP.
FT	DOMAIN	101	167	VWFC.
FT	DOMAIN	198	243	TSP TYPE-1.
FT	DOMAIN	247	349	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	256	330	CTCK.
FT	DISULFID	256	293	BY SIMILARITY.
FT	DISULFID	273	307	BY SIMILARITY.
FT	DISULFID	287	325	BY SIMILARITY.
FT	DISULFID	292	329	BY SIMILARITY.
FT	CONFLICT	28	28	D -> DC (IN REF. 1).
FT	CONFLICT	58	58	MISSING (IN REF. 1).
FT	CONFLICT	88	88	A -> T (IN REF. 1).
FT	CONFLICT	103	104	CV -> YI (IN REF. 1).
FT	CONFLICT	166	169	CDEP -> SRDE (IN REF. 1).
FT	CONFLICT	184	184	P -> L (IN REF. 1).
FT	CONFLICT	200	200	L -> Q (IN REF. 1).
FT	CONFLICT	209	209	C -> Y (IN REF. 1).
FT	CONFLICT	269	269	E -> Q (IN REF. 1).
FT	CONFLICT	284	284	C -> F (IN REF. 1).
SQ	SEQUENCE	349 AA;	37924 MW;	5FFC8E83E8F4F99 CRC64;

Query Match 44.6%; Score 937; DB 1; Length 349;
 Best Local Similarity 44.7%; Pred. No. 8.5e-66;
 Matches 172; Conservative 65; Mismatches 100; Indels 48; Gaps 8;

QY	1	MSSTFRTLAVAVTLHLRLALST	-----CPAACHCPL-EAPKCAPGVGLVDDGCGC	52
Db	1	MSATGLGVRCAFVILL	-----LALCSRPASSQDCSPAGPAPRCFAGVSLVLDGCGC	56
QY	53	CKVCAKQALNEDCSKTQPCDHTKGLCNFGASSTALKGICRAQSEGRPCVNSRIYQNGES	112	
Db	57	CRVCAKQALSELCTERDPCDPKHLGFCDFGSPANRIGVCTAK-DGAPCVFGTGVVQSGES	115	
QY	113	FPQNCKHQCTCDGAVGCIPLCPQBSLPLNLCGNPRLVKVSGQCCWVCDDESIKDSL	172	
Db	116	FQSCCKYQCTCLDGSVGVPLCSVDVRLFPSPDPFPRVKLPFGKCCWVCDDEP	169	
QY	173	DDODDLGLDAGEVELTRNNELIATGKG-SSLUKLPLVFGTEPRVLPNPLHAHQKCIQVOT	231	
Db	170	-----KEHTVVGPAALAVRPEDTGPDTMI-----RANCLVOT	203	
QY	232	TWSQCSKSGTGITRVNDNPECLVKETRICVVRPGQGVVSSLKKGKCKSKTKKSP	291	
Db	204	TEWSACSTCGMGISTRVNDNNAFCRLKQSLCVRPCVCEADLEENIKKGGKCIPTPKIS	263	
QY	292	EPVRTYAGCSVKVYRKYCGSCVDGRCCTPLQRTVKMRFCRCDGEMFSGKNVAMIOSC	351	
Db	264	KPIKELSCSTSMKTYRAKFGVGVCDTGRCCTPHRTTILPVEFKCPDGEVKKSMFIKTC	323	

RESULT 9

CTGF_PIG	STANDARD;	PRT;	349 AA.
ID	CTGF_PIG		
AC	019113;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Connective tissue growth factor precursor.		
GN	CTGF OR CCN2.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Uterus;		
RX	MEDLINE=97390475; PubMed=9242708;		
RA	Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,		
RA	Harding P.A.;		


```

CC EMBL; X59284; CAA41975.1; -
DR PIR; S20078; S20078.
DR InterPro; IPR006208; Cys knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000864; TSPI.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFEB; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSPI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00922; TSPI; 1.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Proto-oncogene; Growth factor; Signal.
FT SIGNAL 1 24
FT CHAIN 25 351
FT DOMAIN 31 103
FT DOMAIN 104 170
FT DOMAIN 201 246
FT DOMAIN 258 332
FT DISULFID 258 295
FT DISULFID 275 309
FT DISULFID 286 325
FT DISULFID 289 327
FT DISULFID 294 331
FT CARBOHYD 274 274
SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA3058C6797 CRC64;

Query Match
Best Local Similarity 40.9%; Score 861; DB 1; Length 351;
Matches 165; Conservative 44; Mismatches 110; Indels 52; Gaps 7;

QY 9 LAVAVTLHLTRLA-----LSTCPAAC--HCPLEAPKAPGVGLVRDGGCGCKVCAQLN 61
Db 9 LPVLLLLLLLLPCVSGREACRPGCGRCFPAEPPRCAPGVPAVLPGCGCLVCARQG 68

QY 62 EDCSTQPCDHTKGLCNFGASSTALKICRAQSEGRPCVNSRIYONGSFQNCCKHQ 121
Db 69 ESCSPLLCDESGGLYCDRGPEDGGAGICMV-LEGDNCVFDGMVYRNGETFPSCYQC 127

QY 122 TCIDGAVGICPLCPQELSLPNLGNPNRLVKVSGCCCEWVCD-EDSI-----KDSLDQ 175
Db 128 TCRDQIGCLPRCNLGLLLPGDPCFPFRKIEVPGCECKWVCDPRDEVLLGGFAMAYRQ 187

QY 176 DLLGLDASEVELTRNNELIAGKSSLSKRLPVFGTEPRVLFNPLHAHQKCIQVTTWS 235
Db 188 EATLIDVSD-----SSANCIEQTTEWS 210

QY 236 QCSKCGTGTSTRTNDNPECLVKETRICVRPCQGVYSSLLKGGKCKSKTKSPFVR 295
Db 211 ACSKSCGMGFSTRTNRNQCEMVKQTRLCNMRPCENEPSPD-KKGKKCIQTKSKMAVR 269

QY 296 FTYACSSVKYRKYCCSGVDGRCCTPLQRTVKMRPCEDGEMFKNVMIOSCKNY 355
Db 270 FEYKNCISVQYKRYCLNDRCCPTPHNTKTIQVEFRCPQGRFLKPKPMLINTVCVHG 329

QY 356 NCPHNEASFR 366
Db 330 NCPQSNNAFFQ 340

RESULT 11
NOV_COTJA
ID NOV_COTJA STANDARD; PRT; 353 AA.

```

```

AC P42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV OR CCN3.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiskirchen R., Bister K.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Immediate-early protein likely to play a role in cell
CC growth regulation (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the CCN family.
CC -!- SIMILARITY: Contains 1 IGFEB domain.
CC -!- SIMILARITY: Contains 1 VWF domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; UI3063; AAA21128.1; -
CC InterPro; IPR006208; Cys_knot.
CC InterPro; IPR006207; Cys_knot_C.
CC InterPro; IPR000867; Insl_gro_fac_pr.
CC InterPro; IPR000864; TSPI.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PF00219; IGFEB; 1.
CC Pfam; PF00090; tsp_1; 1.
CC Pfam; PF00093; vwc; 1.
CC SMART; SM00041; CT; 1.
CC SMART; SM00121; IB; 1.
CC SMART; SM00209; TSPI; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS01225; CTCK_2; 1.
CC PROSITE; PS00222; IGF_BINDING; 1.
CC PROSITE; PS00922; TSPI; 1.
CC PROSITE; PS01208; VWF_1; 1.
CC PROSITE; PS0184; VWF_2; 1.
KW Proto-oncogene; Growth factor; Signal.
FT SIGNAL 1 26
FT CHAIN 27 353
FT DOMAIN 33 105
FT DOMAIN 106 172
FT DOMAIN 203 248
FT DOMAIN 260 334
FT DISULFID 260 297
FT DISULFID 277 311
FT DISULFID 288 327
FT DISULFID 291 329
FT DISULFID 296 333
FT CARBOHYD 276 276
SQ SEQUENCE 353 AA; 38667 MW; 717D9F8533882E89 CRC64;

Query Match
Best Local Similarity 40.7%; Score 856; DB 1; Length 353;
Matches 163; Conservative 47; Mismatches 108; Indels 56; Gaps 7;

QY 10 AVAVTLHLTRLA-----STCPAAC--HCPLEAPKAPGVGLVRDGGCGCKVCAK 58
Db 10 AVAVTLHLTRLA-----STCPAAC--HCPLEAPKAPGVGLVRDGGCGCKVCAK 58

```

```
Db 8 SLPVLLLLLLLLLRSEVNGREAPCPRCGGRCAPPAEPPRCAPGVPAVLGGCCGLVCAR 67
QY 59 QLNEDSKTQPCDHTKGLSCNFGASTALKGTICRAQSEGRPEYNSRIYONGESTQPNCK 118
Db 68 QRGESCPPLPCDSEGLYGRGPDGGGTGICMV-LEGDNCVFDGMIRNGETFPQSCCK 126
QY 119 HQCTCIDGAVGCIPLCPQBSLPLNLGCPNPRLVKVSQCCCEWVCD-EDSI-----KDSL 172
Db 127 YQCTCRDGOIGCLPCNLGLLPGDPPFPKRIEVPGECEKXWCEPRDEVLLGGFAMAA 186
QY 173 DDQDLGLDASEVELTRNNELIAIGKSSLRKLFPVGTPEPRVLPNPLHAHQKQIVQTT 232
Db 187 YROEATLIGDIVSD-----SSANCIQET 209
QY 233 SWQCKSKCTGTISTVTNDNPECLRVKTRICEVRPCQPVYVSLKKGKSKTKKSP 292
Db 210 EWSACRSRSGMSTVTRNQCENVMVQTRLCWRPCENEPSPD-KGKKKIRTKSKMK 268
QY 293 PVRTYAGSSVKYKPKYKCGSVDRGCTCTPTQTRVQMRPCDEGEMPSKVMNIQSCK 352
Db 269 AVRFYKNTKTSVQYKPRYVGLNDRGCTCTPTHTKTIQVEPRCPQCKFLKPEMLINTCV 328
QY 353 CNYNCPHPNEASFR 366
Db 329 CHGNCPOSNAFFQ 342

RESULT 12
NOV_HUMAN
ID NOV_HUMAN STANDARD; PRT; 357 AA.
AC P46745; Q96BY5;
DT 01-FEB-1996 (Rel. 33, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NovH) (Nephroblastoma overexpressed
  gene protein homolog).
GN NOV OR CN3 OR NOVH OR IGFBP3.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94336229; PubMed=7520150;
RA Martinerie C., Huff V., Joubert I., Badzioch M., Saunders G.F.,
RA Strong L.C., Perbal B.;
RT "Structural analysis of the human nov proto-oncogene and expression
  in Wilms tumor."
RL Oncogene 9:2729-2732(1994).
[2]
SEQUENCE FROM N.A.
RA Jiang D., Gou D., Li W.;
RT "Cloning, sequencing and expression of human nov gene."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters H.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
```

```
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skolaka U., Smailus D.E.,
RA Schmerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
INTERACTION WITH FBLN1.
MEDLINE=99128329; PubMed=9927660;
RX Perbal B., Martinerie C., Sainson R., Werner M., He B., Roizman B.;
RA "The C-terminal domain of the regulatory protein NOVH is sufficient to
  promote interaction with fibulin 1C: a clue for a role of NOVH in
  cell-adhesion signaling."
RL Proc. Natl. Acad. Sci. U.S.A. 96:869-874(1999).
CC -!- FUNCTION: Immediate-early protein likely to play a role in cell
  growth regulation (By similarity).
CC -!- SUBUNIT: Interacts with FBLN1.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Increased expression in Wilms tumor of the
  stromal type.
CC -!- SIMILARITY: Belongs to the CN family.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 WFV domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X78351; CAAS5146.1; -
EMBL; X78352; CAAS5146.1; JOINED.
EMBL; X78353; CAAS5146.1; JOINED.
EMBL; X78354; CAAS5146.1; JOINED.
EMBL; X96584; CAAS5146.1; -
EMBL; AY082381; AAL92490.1; -
EMBL; BC015028; AAL15028.1; -
PIR; I38069; I38069.
Gene; HGNC:7885; NOV.
MIM; 164958; -
InterPro; IPR006208; Cys_knot.
InterPro; IPR006207; Cys_knot_C.
InterPro; IPR000867; Ins1_gro_fac_pr.
InterPro; IPR000884; TSP1.
Pfam; PF00007; Cys_knot; 1.
Pfam; PF00219; IGFBP; 1.
Pfam; PF00090; tsp_1; 1.
Pfam; PF00093; vwc; 1.
SMART; SM00041; CT; 1.
SMART; SM00121; IB; 1.
SMART; SM00209; TSP1; 1.
SMART; SM00214; VWC; 1.
PROSITE; PS01185; CTCK_1; 1.
PROSITE; PS01225; CTCK_2; 1.
PROSITE; PS00222; IGF_BINDING; 1.
PROSITE; PS00092; TSP1; 1.
PROSITE; PS01208; WFV; 1.
PROSITE; PS0184; WFV_2; 1.
KW Growth factor; Signal.
FT SIGNAL 1 27
FT CHAIN 28 357
FT DOMAIN 35 107
FT DOMAIN 108 174
FT DOMAIN 205 250
FT DOMAIN 264 338
FT DISULFID 264 301
FT DISULFID 281 315
FT DISULFID 292 331
POTENTIAL.
NOV PROTEIN HOMOLOG.
IGFBP.
WFV.
TSP TYPE-1.
CTCK.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
```


RESULT 14

NOV MOUSE
 ID - NOV MOUSE STANDARD; PRT: 354 AA.
 AC Q64299; Q8CA67;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed gene protein homolog)
 GN NOV OR CCN3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV, and ICR; TISSUE=Brain;
 RX MEDLINE=97131523; PubMed=8975721;
 RA Snaith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinierie C.,
 RA Perbal B., Schofield P.N., Boulter C.A.;
 RT "Genomic structure and chromosomal mapping of the mouse nov gene."
 RL Genomics 38:425-428(1996).
 [2]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=96204003; PubMed=8622864;
 RA Martinierie C., Chevalier G., Rauscher F.J. III, Perbal B.;
 RT "Regulation of nov by Wt1: a potential role for nov in
 RT nephrogenesis."
 RL Oncogene 12:1479-1492 (1996).
 [3]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryonic head, and Spinal cord;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaudo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Regi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Gough J.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume M., Inotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Saeki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino K., Sasaki D., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ; TISSUE=Breast tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Stapleton M., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Brownstein M.J., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Rana S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Grimwood J., Schmutz J., Myers R.M.,
 RA Rodriguez A.C., Krzyzinski M.I., Skalska U., Smalilus D.E.,
 RA Butterfield Y.S.N., Jones S.J.M., Marra M.A.;
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Immediate-early protein likely to play a role in cell
 CC growth regulation (By similarity).
 CC -1- SUBUNIT: Interacts with FBLN1 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the CCN family.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFBC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL; X97863; CA66457.1; -;
 DR EMBL; Y0257; CAA70454.1; -;
 DR EMBL; X96585; CAA65404.1; -;
 DR EMBL; AK039481; BAC30363.1; -;
 DR EMBL; AK081944; BAC38378.1; -;
 DR EMBL; BC003774; AA03774.1; -;
 DR MGI; MGI:109185; Nov. Cys knot.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsep_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS01208; VWF_C_1; 1.
 DR PROSITE; PS01084; VWF_C_2; 1.
 KW Growth factor; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 354
 FT DOMAIN 29 101
 FT DOMAIN 102 168
 FT DOMAIN 202 247
 FT DOMAIN 261 335
 FT DOMAIN 261 298
 FT DISULFID 278 312
 FT DISULFID 289 328
 FT DISULFID 292 330
 FT DISULFID 297 334
 FT CARBOHYD 91
 FT
 FT POTENTIAL.
 FT NOV PROTEIN HOMOLOG.
 FT IGFBP.
 FT VWF_C.
 FT TSP TYPE-1.
 FT CTCK.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT

FT CARBOHYD 277 277 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 6 R -> G (IN REF. 3; BAC30363).
SQ SEQUENCE 354 AA; 38928 MW; 08CEBCFC67829DE CRC64;

Query Match
Best Local Similarity 39.3%; Score 827; DB 1; Length 354;
Matches 161; Conservative 51; Mismatches 114; Indels 38; Gaps 9;

QY 7 RFLAVAVTLLH-LTRALS--TCPAAC--HCPLEAPKAPGVGLVRDGGCGCKVCAKQNLNE 62
DB 8 RCLCLGFLHLLSQVSASLRCPSPKPSISPTCAPGVRSVLDGSCCFVCARQGE 67
QY 63 DCSKTOPCDHTKGLCNFGASSTALKICRAQSEGRPCENRSRYQNGESFQPNCKHOCT 122
DB 58 SCSENRPCDQSSGLCDRSAPNNQGLCW-PEGDNVFGVIYRNGEKEPNCQVCT 126
QY 123 CIDGAVGICPICQBELSLPNLGNPNRLVKSQCCEWVC--DEDSIKDLDODDILLG 180
DB 127 CRDQIGICLPRQLDVLFLFGPDCAFRKVAVFGECCKWTCGSDQGTQGLG---LA 182
QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAGQKCIQVOTTSQCSKS 240
DB 183 LPAYRPEATVGEV-----SDSSI-----NCLEQTTEWSACS 216
QY 241 COTGISTRVNDNPECLVKEIRICEVRPCQ--PVYSLKXKCKSKTKSPVPRFYA 299
DB 217 CGMGVSTRVNRNRCQENVKQTRLCIVRPECQEPPEVTDKKGKCLRTKSLKAHLQFE 276
QY 300 GCSVVKYRKYKCGSVGRCTPLQTVKMRFCEDGEMFSGNVMQISCKNVCNCPH 359
DB 277 NCTSLYTYRPRFCGVCSGDRCTPHNTKTIQVBPQCLFEBIIRKVPVWVGICTYSNCPQ 336
QY 360 PNEA 363
DB 337 NNEA 340

RESULT 15
NOV_XENLA STANDARD; PRT; 343 AA.
AC P51609;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (Xnov).
GN NOV OR CCN3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96257227; PubMed=8656280;
RA Ying Z., King M.L.;
RT "Isolation and characterization of xnov, a Xenopus laevis ortholog of
the chicken nov gene";
RL Gene 171:243-248(1996).
CC -!- FUNCTION: Immediate-early protein likely to play a role in cell
growth regulation (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the CCN family.
CC -!- SIMILARITY: Belongs to the IGFBP domain.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
CC EMBL; U37063; AAB17036.1; -
CC InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot.C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF.C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwg; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWF; 1.
DR PROSITE; PS01185; CTCK_1; FALSE NEG.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1; 1.
DR PROSITE; PS50184; VWF_2; 1.
KW Growth factor; Signal.
FT SIGNAL 1 18
FT CHAIN 19 343 NOV PROTEIN HOMOLOG.
FT DOMAIN 21 92 IGFBP.
FT DOMAIN 93 159 VWF.
FT DOMAIN 190 235 TSP TYPE-1.
FT DOMAIN 249 323 CTCK.
FT DISULFID 249 286 BY SIMILARITY.
FT DISULFID 266 300 BY SIMILARITY.
FT DISULFID 277 316 BY SIMILARITY.
FT DISULFID 280 318 BY SIMILARITY.
FT DISULFID 285 322 BY SIMILARITY.
FT CARBOHYD 265 265 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 343 AA; 38070 MW; 677D7078EB21365F CRC64;

Query Match 39.2%; Score 824.5; DB 1; Length 343;
Best Local Similarity 42.1%; Pred. No. 4.6e-57;
Matches 152; Conservative 49; Mismatches 107; Indels 53; Gaps 7;

QY 9 LAVAVTLLHTRALS--TCPAAC--HCPLEAPKAPGVGLVRDGGCGCKVCAKQNLNEDCSKT 67
DB 5 LALCFILL-IQVASKQCPSCDQCPEEPSCAPSVLLILDGCGCCFVCARQEGESCSSL 63
QY 68 QPCDHTKGLCNFGASSTALKICRAQSEGRPCENRSRYQNGESFQPNCKHOCTCIDGA 127
DB 64 NPCQEDKGLYCEFNADPRMETGTCTMA-LEGNSCVFDGVVYRNRSEFPSCYKHTCLNGH 122
QY 128 VGCIPCLCPQELSLPNLGNPNRLVKSQCCEWVCDEDSIKDLD-----DQDDL 178
DB 123 IGVPRCNLDLLLPDPCPPFRVYKVPFGCCCKWVDS---KEEMAIGGFAMAAYRPEAT 179
QY 179 LGDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAGQKCIQVOTTSQCS 238
DB 180 LGIDASDTSFA-----CIAQTTEWSACS 202
QY 239 KSCGTGISTRVNDNPECLVKEIRICEVRPC--GQPVYSLKXKCKSKTKSPVPRFT 297
DB 203 KTCMGVSTRVNRNRCQENVKQTRLCIVRPECQEPPEVTDKKGKCLRTKSLKAHLQFE 262
QY 298 YAGCSSVKYRKYKCGSVGRCTPLQTVKMRFCEDGEMFSGNVMQISCKNVCNCPH 357
DB 263 YKNTSVQPKYKPKFCQCSGDRCTPHNTKTIQVBPQCLFEBIIRKVPVWVGICTYCNCP 322
QY 358 P 358
DB 323 P 323

Search completed: April 22, 2004, 17:57:04
Job time : 11.4724 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 17:53:04 ; Search time 38.3987 Seconds
(without alignments)
3114.206 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103

Sequence: 1 MSSSTPTLAVATLLHLTR.....PNEASRLYSLFNDIHKRD 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	98.7	379	11 Q9VTM9	Q9VTM9 rattus norv
2	1624.5	77.2	375	13 Q98TK5	Q98TK5 xenopus lae
3	1455	69.2	334	4 Q9UID7	Q9UID7 homo sapien
4	938	44.6	344	13 Q98TQ8	Q98TQ8 gallus gall
5	937	44.6	343	13 Q42607	Q42607 xenopus lae
6	920.5	43.8	347	13 Q9PT80	Q9PT80 notophthalm
7	891	42.4	349	6 Q97765	Q97765 sus scrofa
8	635.5	30.2	331	4 Q95958	Q95958 homo sapien
9	495.5	23.6	176	13 Q9FS86	Q9FS86 gallus gall
10	374.5	17.8	128	11 Q9R2C0	Q9R2C0 rattus norv
11	352	16.7	125	6 Q862T0	Q862T0 bos taurus
12	345	16.4	113	11 Q91V29	Q91V29 rattus norv
13	343	16.3	119	11 Q91V29	Q91V29 mus musculu
14	343	16.3	119	11 Q920W6	Q920W6 mus spicile
15	341	16.2	100	4 Q9UDL6	Q9UDL6 homo sapien
16	298.5	14.2	374	5 Q9VVK0	Q9VVK0 drosophila

17	296	14.1	470	5 Q9VVK3	Q9VVK3 drosophila
18	293	13.9	230	4 Q8WYK7	Q8WYK7 homo sapien
19	195	9.3	77	4 Q9UDE4	Q9UDE4 homo sapien
20	184	8.7	70	13 Q9DFZ1	Q9DFZ1 scyllorhinu
21	179	8.5	2327	13 Q9IBG7	Q9IBG7 xenopus lae
22	173.5	8.3	4007	4 Q86XX4	Q86XX4 homo sapien
23	171	8.1	1045	5 Q8T3A6	Q8T3A6 caenorhabdi
24	171	8.1	1070	5 Q8T3A7	Q8T3A7 caenorhabdi
25	171	8.1	1111	5 Q8XWD6	Q8XWD6 caenorhabdi
26	167.5	8.0	792	13 Q90Z43	Q90Z43 gallus gall
27	167.5	8.0	1095	13 Q90XG4	Q90XG4 gallus gall
28	167	7.9	58	6 Q97574	Q97574 bos taurus
29	166	7.9	1034	11 Q35888	Q35888 rattus norv
30	165.5	7.9	950	13 Q90Z44	Q90Z44 gallus gall
31	163	7.8	1036	4 Q9NZV1	Q9NZV1 homo sapien
32	162	7.7	360	5 Q86AK7	Q86AK7 dictyosteli
33	161.5	7.7	406	11 Q92513	Q92513 mus musculu
34	161.5	7.7	426	11 Q8VEA6	Q8VEA6 mus musculu
35	160.5	7.6	2104	5 Q964N4	Q964N4 caenorhabdi
36	160.5	7.6	2104	5 Q91281	Q91281 caenorhabdi
37	158.5	7.5	1028	11 Q9JLL0	Q9JLL0 mus musculu
38	158.5	7.5	4998	11 Q8CG65	Q8CG65 mus musculu
39	158	7.5	1664	5 Q9TVQ2	Q9TVQ2 caenorhabdi
40	157.5	7.5	13288	6 Q18758	Q18758 sus scrofa
41	157	7.5	894	11 Q88715	Q88715 mus musculu
42	157	7.5	1350	11 Q80Z20	Q80Z20 mus musculu
43	156.5	7.4	1574	11 Q88281	Q88281 rattus norv
44	156.5	7.4	1698	5 Q94438	Q94438 chironomus
45	156	7.4	1025	11 Q9Z166	Q9Z166 mus musculu

ALIGNMENTS

RESULT 1

Q9VTM9 PRELIMINARY; PRT; 379 AA.

AC Q9VTM9; 01-NOV-1999 (TREMELrel. 12, Created)

DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE CYR61 precursor.

GN CYR61.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Izm; TISSUE=Aorta;

RA Unoki H., Yonekura H., Furukawa K., Yamamoto H.;

RT "Rat Cy-61 mRNA."

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB015877; BAA78339.1; -.

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0005520; F:insulin-like growth factor binding; IEA.

DR GO; GO:0001558; P:regulation of cell growth; IEA.

DR InterPro; IPR006208; Cys_knot.

DR InterPro; IPR006207; Cys_knot_C.

DR InterPro; IPR000887; Insi_gro_fac_pr.

DR InterPro; IPR000884; TSPL.

DR Pfam; PF00007; Cys_knot; 1.

DR Pfam; PF00219; IGFBP; 1.

DR Pfam; PF00090; tsp_1; 1.

DR Pfam; PF00093; wgc_1; 1.

DR SMART; SM00041; VEG; 1.

DR SMART; SM00221; IB; 1.

DR SMART; SM00209; TSPL; 1.

DR SMART; SM00214; VMC; 1.

DR PROSITE; PS01185; CTCK 1; 1.

DR PROSITE; PS01225; CTCK 2; 1.

DR PROSITE; PS00222; IGF_BINDING; 1.


```

RESULT 12
Q92164 PRELIMINARY; PRT; 113 AA.
AC Q92164;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Connective tissue growth factor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99150373; PubMed=10026205;
RA Murphy M., Godson C., Cannon S., Kato S., Mackenzie H.S., Martin F.,
RA Brady H.R.;
RT "Suppression subtractive hybridization identifies high glucose levels
RT as a stimulus for expression of connective tissue growth factor and
RT other genes in human mesangial cells.";
RL J. Biol. Chem. 274:5830-5834(1999).
DR EMBL; AF079531; AB02838.1; -
DR InterPro; IPR006208; Cys_knot.
DR Pfam; PF00007; Cys_knot; 1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12767 MW; 23EA69FC0A0635E CRC64;

Query Match 16.4%; Score 345; DB 11; Length 113;
Best Local Similarity 55.1%; Pred. No. 4.8e-25;
Matches 59; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 245 ISTRVNDNPECHLVKTRICEVRPGQPVYSLKGGKCKTKKSPFVRYTYACSSV 304
DB 1 ISTRVNDNFTFCRLKQSLRCLWVRPCEADLEENIKGKCIPTPKIAKPVKFSCTSV 60

QY 305 KTVRPKYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFSKNVMQISC 351
DB 61 KTVRAKFCGVCTDGRCTPHRTTLTPVEFKCPHGEIMKKNMFIKTC 107

RESULT 13
Q91V29 PRELIMINARY; PRT; 119 AA.
AC Q91V29;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE FISP-12 protein (Fragment).
GN FISP-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=various strains;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Kobayakawa H., Saitou N.;
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=pgn2;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039094; BAB68618.1; -
DR EMBL; AB039095; BAB68619.1; -
DR EMBL; AB039096; BAB68620.1; -
DR EMBL; AB039097; BAB68621.1; -
DR EMBL; AB039098; BAB68622.1; -
DR EMBL; AB039099; BAB68623.1; -
DR EMBL; AB039100; BAB68624.1; -
DR EMBL; AB039101; BAB68625.1; -
DR EMBL; AB039102; BAB68626.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13626 MW; C7C8AD253DF331CE CRC64;

Query Match 16.3%; Score 343; DB 11; Length 119;
Best Local Similarity 48.7%; Pred. No. 7.9e-25;
Matches 57; Conservative 25; Mismatches 33; Indels 2; Gaps 1;

QY 260 KETRICVRPGQPVYSLKGGKCKTKKSPFVRYTYACSSVKKYRPKYCGSCVDGR 319
DB 2 KQSLRCLWVRPCEADLEENIKGKCIPTPKIAKPVKFSCTSVKTYRAKFCGVCTDGR 61

QY 320 CCTPLQTRTVKMRFCEDGEMFSKNVMQISCCKNYNCPHNEASFLY--SLFNDI 374
DB 62 CCTPHRTTLTPVEFKCPDGEIMKKNMFIKTCACHYNCPGDNDIFESLYYKMYGDM 118

RESULT 14
Q920W6 PRELIMINARY; PRT; 119 AA.
AC Q920W6;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE FISP-12 protein (Fragment).
GN FISP-12.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZEN.
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039103; BAB68627.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR Pfam; PF00007; Cys_knot; 1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13626 MW; C7C8AD253DF331CE CRC64;

Query Match 16.3%; Score 343; DB 11; Length 119;
Best Local Similarity 48.7%; Pred. No. 7.9e-25;
Matches 57; Conservative 25; Mismatches 33; Indels 2; Gaps 1;

QY 260 KETRICVRPGQPVYSLKGGKCKTKKSPFVRYTYACSSVKKYRPKYCGSCVDGR 319
DB 2 KQSLRCLWVRPCEADLEENIKGKCIPTPKIAKPVKFSCTSVKTYRAKFCGVCTDGR 61

```


